Genetics of Influenza Viruses

Edited by
Peter Palese and David W. Kingsbury

Springer-Verlag Wien New York

Professor Dr. *Peter Palese*
Mount Sinai School of Medicine, The Mount Sinai Hospital, New York, N.Y., U.S.A.

Professor Dr. *David W. Kingsbury*
St. Jude Children's Research Hospital, Memphis, Tenn., U.S.A.

With 62 partly coloured Figures

Library of Congress Cataloging in Publication Data. Main entry under title: Genetics of influenza viruses. Includes bibliographical references and indexes. 1. Influenza viruses—Genetics. I. Palese, Peter, 1944—. II. Kingsbury, David W. QR201.I6G47. 1983. 616.2'030194. 83-16680

ISBN-13:978-3-7091-8708-1 e-ISBN-13:978-3-7091-8706-7
DOI: 10.1007/978-3-7091-8706-7

Preface

With the advent of genetic engineering methods and improved biochemical techniques, much has been learned about the replication of influenza viruses, their structure and their epidemiology. It appears that the time is ripe to review these efforts and to provide a molecular perspective of influenza virology. It is hoped that this book will stimulate our thinking, help us in designing new experiments, and possibly show avenues leading to the control of the diseases associated with influenza viruses.

Peter Palese, New York, N.Y.
August 1983 *David W. Kingsbury,* Memphis, Tenn.

Contents

List of Contributors

Dr. *G. M. Air,* Department of Microbiology, University of Alabama Medical Center, 520 11th Street South, University Station, Birmingham, AL 35294, U.S.A.

Dr. *R. W. Compans,* Department of Microbiology, University of Alabama Medical Center, 520 11th Street South, University Station, Birmingham, AL 35294, U.S.A.

Dr. *Mary-Jane Gething,* Cold Spring Harbor Laboratory, P.O.Box 100, Cold Spring Harbor, NY 11724, U.S.A.

Dr. *E. D. Kilbourne,* Department of Microbiology, Mount Sinai School of Medicine, 5th Avenue and 100th Street, New York, NY 10029, U.S.A.

Dr. *D. W. Kingsbury,* St. Jude Children's Research Hospital, 332 North Lauderdale, P.O.Box 318, Memphis, TN 38101, U.S.A.

Dr. *R. M. Krug,* Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, U.S.A.

Dr. *R. A. Lamb,* Department of Biochemistry, Molecular and Cell Biology, Northwestern University, 2153 Sheridan Road, Evanston, IL 60201, U.S.A.

Dr. *W. G. Laver,* John Curtin Medical School, Australian National University, P.O.Box 334, Canberra City, A.C.T. 2601, Australia.

Dr. *B. W. J. Mahy,* Division of Virology, Department of Pathology, University of Cambridge, Laboratories Block, Addenbrookes Hospital, Hills Road, Cambridge CB 2 2QQ, United Kingdom.

Dr. *D. P. Nayak,* Department of Microbiology and Immunology, School of Medicine, University of California, Los Angeles, CA 90024, U.S.A.

Dr. *P. Palese,* Mount Sinai School of Medicine, 5th Avenue and 100th Street, New York, NY 10029, U.S.A.

Dr. *J. Sambrook,* Cold Spring Harbor Laboratory, P.O.Box 100, Cold Spring Harbor, NY 11724, U.S.A.

Dr. *C. Scholtissek,* Justus Liebig-Universität, Frankfurter Strasse 107, D-6300 Giessen, Federal Republic of Germany.

Dr. *J. L. Schulman,* Department of Microbiology, Mount Sinai School of Medicine, 5th Avenue and 100th Street, New York, NY 10029, U.S.A.

Dr. *N. Sivasubramanian,* Department of Microbiology and Immunology, School of Medicine, University of California, Los Angeles, CA 90024, U.S.A.

Dr. *R. G. Webster,* Division of Virology and Molecular Biology, St. Jude Children's Research Hospital, 332 North Lauderdale, P.O.Box 318, Memphis, TN 38101, U.S.A.

Dr. *J. F. Young,* Mount Sinai School of Medicine, 5th Avenue and 100th Street, New York, NY 10029, U.S.A.

1

The Evolution of Influenza Viral Genetics—A Perspective

E. D. Kilbourne

I. Introduction

It is probable that the course of human history has been affected many times by mutations in the hemagglutinin protein of influenza A viruses. Because the clinical and epidemiological features of past and modern epidemics are so similar, we can assume that past epidemics were caused by antigenically variable viruses like contemporary strains now under study. The legitimacy of this assumption is strengthened by serologic, and more recently, molecular genetic evidence of the return or recycling of virus variants from the past. By inference, contemporary studies of the genetics of influenza viruses have not only archaeological but predictive implications for this recycling virus.

While variants of other human viruses have evolved into multiple contemporaneous species, the influenza A viruses—although notable for their variation—appear to coexist in man with difficulty, so that only one or two major variants evolve to produce disease in any one year, and such "subtypes" survive but one to three decades. In this accelerated evolution of a virus doomed to prompt disappearance as immunity in the population rises lies the fascination of influenza and the problem of its prevention. As in no other disease, the genetics of the virus permeates understanding of its pathogenesis and epidemiology, as succeeding chapters in this book will demonstrate.

Influenza virus genetics originated less as a formal genetic system than as an attempt to define and understand the practical problems presented by the ceaseless antigenic variation of a virus capable both of pandemic eruptions and smoldering epidemicity. To be explained as well is the paradox of an unvarying disease caused by a varying virus (Kilbourne, 1980 a, b), an enigma now yielding to the identification of conserved amino acid sequences and variable regions in key structural proteins of the virus (Air, 1981; Jou *et al.*, 1980; Desselberger, Zamecnik and Palese, 1980; Desselberger *et al.*, 1980).

As viruses have been more intensively studied with continually improving biochemical and immunological techniques, the apparent uniqueness of influenza

virus antigenic variation and its seeming hypermutability might be called into question. However, despite the recent evidence that rabies (Koprowski *et al.*, 1980), poliovirus (Nottay *et al.*, 1981) and other RNA viruses (reviewed by Holland *et al.*, 1982) undergo significant antigenic variation, only with influenza does the continual emergence of mutants clearly influence the epidemiology of the disease.

II. The Development of Modern Influenza Viral Genetics

A. Early Evidence of Genetic Variation in the Laboratory

Soon after the initial isolation of the influenza A viruses, serial passage of virus in experimental animals was found to yield virus of enhanced virulence for these laboratory hosts. Such "adaptation", commonly effected in the early days of virology when virus was propagated only in intact animals, obviously represents the selection of viral genetic variants better equipped for replication in the new and alien host. Some years later, the pleiotropic nature of some such mutations was recognized when "adapted" virus was found to react differently with hemagglutination inhibitors than did the original strain (Briody *et al.*, 1955). Conversely, it was shown that virulence for the mouse lung of strains never passaged in mice might develop during chick embryo allantoic passage with mucoprotein inhibitor (Medill-Brown and Briody, 1955). Even earlier, Burnet and Bull (1943) had reported the emergence on chick embryo passage of "D" or derived virus capable of multiplication in the allantoic sac and changed from the original (O) form of virus restricted to amniotic multiplication. Pleiotropically, the D form caused lung lesions in mice and agglutinated chicken RBC with increased efficiency.

In retrospect, these early observations suggest the importance of hemagglutinin-receptor interactions in determining virulence, and clearly illustrate the genetic heterogeneity of influenza viruses as handled in the laboratory. Formal proof of this fact was first presented by Isaacs and Edney (1950) who employed the limiting dilution technique to clone laboratory strains. This technique, "designed to give reasonable statistical assurance that infection had been initiated by a single virus particle (was) the somewhat shaky cornerstone of genetic studies of influenza virus" (Kilbourne, 1963) prior to the development of plaquing systems. Previous studies of viral variation had been conducted with uncloned viruses.

Antigenic variation of field isolates was recognized within a few years after discovery of the virus. Among the early H1 strains, WS and PR 8, the classical initial isolates of influenza A virus from the old and new worlds were found to be antigenically distinguishable from, but related to, Shope's swine virus, isolated earlier in 1931 (reviewed by Hoyle, 1968). Major and globally significant antigenic variation was first encountered in 1946/47 with the advent of the so-called "A prime" strains, which recently reappeared in 1977.

B. Application of Formal Genetic Techniques to Studies of Influenza Virus

It is truly remarkable that during the period 1949–1971, which antedated the devolopment of adequate plaquing systems for influenza viruses, all of the fundamentally important phenomena of influenza virus genetics were described. This advance was dependent to a considerable degree on 1) successful cultivation of the virus in the chick embryo (Smith, 1935; Burnet 1940) and 2) the discovery of viral hemagglutination (Hirst, 1941; McClelland and Hare, 1941). Although not equivalent to cell culture in providing quantifiable numbers of genetically definable cells in an *in vitro* system, the allantoic sac of the chick embryo did provide a single layer of endodermal cells capable of producing 500 to 1400 virus particles per cell (Horsfall, 1955). The allantoic membrane was susceptible to most influenza viruses regardless of their past passage history. Hemagglutination provided a system not available to students of bacterial virus genetics which could measure virus particles as physical or biologic units *in vitro* and in which viral characteristics not dependent on replication for their expression could be defined. The more satisfactory genetic markers used in the occasionally murky early period of influenza virus genetics were defineable *in vitro* directly or indirectly through the hemagglutination reaction.

But more important than these somewhat limited technological advances was the intellectual orientation of investigators such as Burnet, Hirst, Henle and Horsfall, who were aware of the ferment in microbial genetics and quantitative biology, and who sought to find comparable phenomena with this easily handled animal virus. Thus, "recombination" (now more properly defined as genetic reassortment) was discovered by Burnet and Lind (1949) during studies of the interference phenomenon, a phenomenon first remarked with animal viruses in intact animals. Two years later, Henle and Liu (1951) described multiplicity reactivation of both influenza A and B viruses as a phenomenon analogous to multiplicity reactivation with T-even bacteriophage (reviewed by Suguira, 1975) in which complementation among partially inactivated virus particles of the same strain occurs. This phenomenon, initially questioned, was confirmed by studies employing single-cycle infection (Barry, 1961 a, b) or addition of cortisone (Kilbourne, 1955), which greatly amplified reactivation for reasons still unexplained but possibly related to suppressive effects on viral autointerference (Kilbourne, 1957).

C. Genetic Markers

The ambiguity of many early genetic experiments with influenza viruses (and some later ones as well) was due to the unavoidable use of imperfect genetic markers. As animal virology emerged from its original close linkage with pathology, reliance was placed on markers related to pathogenicity of a virus for experimental animals. The polygenic nature of such end points as mouse lung consolidation was recognized early by Burnet, the prime mover in the field. A series of mutational steps appeared necessary in the adaptation of virus to mouse lung replication and for the

production of gross pulmonary lesions (Burnet and Lind, 1954 b). The most useful of the virulence markers has been the neurovirulence manifested by two independently derived mutants of the original WS strain, NWS (Stuart-Harris, 1939) and WS-N (Francis and Moore, 1940). But this marker also is polygenic (Kilbourne, 1963; and see discussion of virulence, below).

All markers dependent for their expression on viral replication, including plaque size and morphology, are subject to phenotypic variation. The best of the replication-dependent markers are conditional in their expression and have tight restriction barriers (as exemplified by temperature sensitive *(ts)* cold adapted *(ca)*, and host range mutants).

As with few other viruses, the influenza viruses bear markers readily demonstrable *in vitro* based on the biological activities of their two surface glycoproteins, the hemagglutinin and neuraminidase. Furthermore, antigenicity of the hemagglutinin has been the most constant and most useful of all markers despite its variability in nature. Minor antigenic variation does not compromise experiments conducted with viruses of different hemagglutinin subtype. Combined with hemagglutinin serotype, virus morphology (whether spherical or filamentous) proved to be a striking and convincing marker of genetic recombination (Kilbourne and Murphy, 1960). It must be said that early work on influenza virus genetics was viewed without conviction by classical microbial geneticists, who were unprepared for the high frequency recombination now known to be characteristic of segmented-genome viruses, and who placed little reliance on the complex viral markers then widely used. It is paradoxical that the influenza viruses were the first animal viruses for which unequivocal recombination was shown, despite the absence initially of plaquing systems and efficient methods for the cloning and selection of recombinant progeny. No doubt the high frequency of genetic reassortiment characteristic of influenza virus permitted the recognition of "recombination" even under these conditions. It is to the credit of both Burnet and Hirst that they recognized that this anomalous rate of genetic interaction was not logically explicable as a crossing-over reaction and required an unconventional explanation. Burnet referred to "redistribution of virulence genes" and the possibility of the genome comprising "a variable number of genetic determinants" (Burnet, 1956). Hirst postulated an "exchange of large pieces" (of the genome) in recombination occurring at a high rate (Hirst, 1962).

D. Development of Plaquing Systems

Development of a plaquing system for influenza viruses in chick embryo fibroblasts (CEF) cells (Staiger, 1964; Simpson and Hirst, 1961) opened the door for the identification of one-step mutants and the estimation of reversion frequency (Staiger, 1964). It also led to confirmation of earlier studies of cross reactivation of UV-inactivated virus (Burnet and Lind, 1954 a; Baron and Jensen, 1955; Gotlieb and Hirst, 1956; Kilbourne and Murphy, 1960) and, most importantly, establishment of a cloning system for certain plaque-forming strains of the virus (Simpson and Hirst, 1961). Later, the use of a continuous human cell line extended the range of viruses in

which plaquing could be carried out and plaque-type recombinants produced (Sugiura and Kilbourne, 1965, 1966). Not until years later was it appreciated that the puzzling inefficiency of influenza viruses as plaque formers was based in large part on the requirement that the viral hemagglutinin must be cleaved by endogenous host cell proteases in order for infection and replication to occur (Lazarowitz and Choppin, 1975; Klenk *et al.*, 1975). The restriction of study to those few viruses (principally the WSN and NWS variants of the original WS strain) which produced distinct plaques greatly curtailed the conduct of genetic experiments. Although it is now known that plaque formation by WSN virus in at least one cell system (MCBK) is dependent upon the WSN neuraminidase (Schulman and Palese, 1977), in general, the capacity of a virus to form plaques is polygenic and phenotypically influenced (as by the action of cellular or exogenous proteases). Nevertheless, influenza virus genetics certainly "came of age" with development of plaquing methods.

E. The Use of Conditional Lethal Mutants

Despite the high frequency recombination of influenza viruses observed during mixed infection and cross-reactivation experiments, the development of a genetic map of the influenza virus genome had been handicapped for lack of single-step mutants bearing good selection markers. The success of Burge and Pfefferkorn (1966) with temperature sensitive (ts) mutants of Sindbis virus prompted Simpson and Hirst (1968) to isolate similar mutants of influenza virus for use in recombination and complementation studies. Earlier, Fenner had emphasized the general promise offered by conditional lethal mutants in the study of animal virus genetics (Fenner, 1965).

Concurrent with the establishment of seven recombination-complementation groups for influenza A viruses in different virus-cell systems (Simpson and Hirst, 1968; Hirst, 1973; Sugiura *et al.*, 1972, 1975; Spring *et al.*, 1975), physico-chemical evidence accrued that influenza virion RNA existed in segments (Duesberg, 1968; Pons and Hirst, 1968). The not unreasonable assumption grew that each RNA piece represented a source of monocistronic message for each putative viral protein. In any case, both genetic and biophysical evidence made it clear that mapping of the genome by the usual genetic approach of measuring recombinational frequencies between mutants would be unrewarding with a virus for which intracistronic recombination had not been demonstrated. A new kind of mapping was forthcoming (see below) which would greatly facilitate influenza virus research. This is not to say that the conventional genetic approach had no further contributions to make. Host range mutants proved valuable in marker rescue studies (Sugiura and Ueda, 1971). Three-factor crosses of two ts mutants demonstrated no linkage between markers and high (40–42.5%) recombination frequency (Nakajima and Sugiura, 1977). More important, these experiments ruled out a significant role for partial heterozygotes in the formation of plaque-forming units, and were consistent with a non-random packaging mechanism for the assembling of viral genes. However, some evidence for the existence of a partially heterozygote virus has been pre-

sented (Scholtissek *et al.*, 1978). Genetic evidence for intracistronic recombination involving the nucleocapside (NP) protein has been supplemented by data from biochemical analysis (Rohde and Scholtissek, 1980). Intrasegmental complementation has been reported (Massicot *et al.*, 1982) as well as extragenic suppression of avirulence, a finding inimical to live virus vaccine development (Murphy *et al.*, 1980).

F. New Approaches in Influenza Virus Genetics

1. *The Biochemical Identification of Viral Gene Products in the Unambiguous Definition of Viral Inheritance*

In all genetic systems the use of markers dependent for their expression on virus replication (including plaque morphology) yields ambiguous results subject to environmental or phenotypic variation. For this reason, heavy reliance was placed in early studies of influenza virus genetics on properties of the virus demonstrable *in vitro,* such as heat stability or (principally) antigenic reactivity. Probably because the antigenic marker was usually defined by the relatively non-specific hemagglutination inhibition test, and because the separability of hemagglutinin and neuraminidase activity was not then appreciated, the antigenic marker was considered to be a single property. However, Hirst and Gotlieb (1953), isolated a stable recombinant "X-3 virus" cross-reactive with both parental virus antisera in HI and neutralization tests. Independently, Kilbourne and Schulman (1965) and Tumova and Pereira (1965) isolated unequivocal antigenic hybrids from different parental virus crosses involving viruses of different subtypes. One of these recombinants, X-7, identified in HI tests as like the NWS parent, was intensively studied and shown by strain-specific complementation tests, plaque size reduction and mouse immunization to contain antigen from the H2N2 virus RI/5$^+$ (Kilbourne and Schulman, 1965; Kilbourne *et al.*, 1966; Jahiel and Kilbourne, 1966). It was also noted that X-7 possessed the high neuraminidase activity characteristic of the H2N2 parent. At that time, Laver had just published on the separation of influenza viral structural proteins (including neuraminidase) by electrophoresis on cellulose acetate (Laver, 1963) and I sought his collaboration in identifying the H2N2 antigenic component. We were able to show that indeed X-7 contained the NWS hemagglutinin and the RI5$^+$ neuraminidase, and by peptide mapping that the NP protein had been acquired from NWS. Thus, even by these relatively simple techniques and as previously noted by Laver (Laver, 1964), differences in the internal viral proteins were found between virus strains (Laver and Kilbourne, 1966). These observations were important for the unequivocal identification by biochemical methods of the origin of gene products contained in a putative recombinant virus. With improvements in the electrophoretic separation of proteins on polyacrylamide electrophoresis (PAGE) resolution of seven virion and (initially) one non-structural protein became possible (Kilbourne *et al.*, 1972). When eight RNA's also were revealed on PAGE it appeared that each RNA segment was a gene from which a single mRNA was transcribed for translation of a single viral protein.

2. *Mapping of the Influenza Virus Genome by Correlative Physico-Chemical and Biological Techniques*

In comparing two human influenza A viruses of different subtype, Palese and Schulman made the important discovery that seven of the eight RNA's of these two viruses (A/PR8/34 H1N1 and A/Hong Kong8/68 H3N2) migrated at different rates on PAGE (Palese and Schulman, 1976). When recombinants of these viruses were prepared, selection with antisera specifically inhibitory to one or the other hemagglutinin or neuraminidase antigen permitted the isolation of viruses antigenically reciprocal (PR8-HK and HK-PR8) with respect to hemagglutinin and neuraminidase. (UV irradiation of the hemagglutinin donor prior to mixed infection favored the selection of asymmetrical reassortants containing all but the hemagglutinin gene from the nonirradiated parent.) Therefore, when gel patterns of parental and antigenically hybrid progeny viruses were compared, it could be shown that only the fourth largest RNA of HK-PR8 differed from that of the PR8 parent but corresponded to that of the HK parent. The converse was true for the PR8-HK reassortant in comparison with Hong Kong virus. Further, the molecular weight estimated for the fourth RNA segment was consistent with the coding capacity for a protein of 70,000 to 80,000 molecular weight, the size of the hemagglutinin protein. Confirmatory evidence for the correspondence of the fourth RNA with the antigenically identified hemagglutinins was provided by similar analysis of back recombinants identical in hemagglutinin-neuraminidase serotype with parental virus (*i.e.*, not antigenically hybrid but reassociated with the original neuraminidases). From these back recombinants which contained only two RNA's derived from the parent of identical serotype, the location of the neuraminidase gene for each virus could be deduced.

These experiments did more than establish a system with which the location of other viral genes could be mapped. They confirmed the truly reassortant nature of influenza virus recombination with the potential for an asymmetrical parental contribution of genes, a point made more systematically in another study (Schulman and Palese, 1976) in which irradiation of one parental virus was also employed to reduce the genetic contribution of that parent. Earlier, the selective effect of inactivation had been empirically utilized in the selection of high yielding recombinant viruses with (Kilbourne *et al.*, 1971) or without (Kilbourne and Murphy, 1960) the use of selective antiserum.

Another significant point made by the Palese-Schulman study was a lack of exact correspondence between RNA size as estimated by gel position and the putative protein for which the mRNA might code. Thus, RNA 5 of the Hong Kong virus and RNA 6 of PR8 were identified as neuraminidase genes. In the later resolution of the genes for the internal proteins, which were not subject to antigenic identification but were identified on protein gels, another exception to the direct correspondence of RNA and protein was encountered. RNA 1 was identified as the gene for P3 (PB2) protein while RNA 2 was the gene for P1 (PB1) and RNA 3 for P2 (PA). The mapping of the genes for proteins not identified by biological activity was dependent upon the development of a gradient gel system in which characteristic virus-specific differences in migration of most viral proteins could be demonstrated

(Ritchey, Palese, and Schulman, 1977). Previously, with influenza B viruses on conventional gels, strain-specific differences in the migration only of hemagglutinin proteins had been found (Tobita and Kilbourne, 1975).

In completion of the map, inferences based on RNA-protein size were necessary, but confirmation has come not only from *in vitro* translation of mRNA (Ritchey and Palese, 1976) but also from oligonucleotide mapping of isolated RNA segments (Young *et al.*, 1981) or by peptide mapping of proteins or by RNA sequencing (reviewed by Webster *et al.*, 1982) in which sufficient conservation of structure permits the ready identification of a gene or gene product as "M, NP", or "NS". The ultimate proof, of course, has come from gene cloning and expression (see Chapter 6).

3. *The Application of Molecular Biological Techniques to the Study of Viral Genetic Variation*

Laver and Webster (1968) pioneered in the application of peptide mapping to study of the viral glycoproteins of major and minor antigenic variants of influenza viruses. Indeed, on the basis of these studies, the magnitude of difference discernible by antigenic analysis was confirmed and the fundamentally different nature of minor (drift) and major (shift) antigenic variation was pointed out (Webster and Laver, 1975). Furthermore, their approach revealed differences in internal proteins (Webster *et al.*, 1982), suggesting that variation was not restricted merely to the surface antigens.

4. *Oligonucleotide Mapping of Viral RNA's*

Like peptide mapping of the gene products, oligonucleotide mapping of the viral genes provided a convenient method of detecting similarities and dissimilarities (mutations) between viruses, particularly when correlated with established RNA sequences (Young *et al.*, 1981). The technique is best suited for discrimination among closely related genes with overall base homologies of over 90%. This method has contributed fundamentally to the analysis of viral variation and recombination in nature.

5. *Contribution of Protein and RNA Sequencing to Influenza Viral Genetics—Intragenic Mapping*

It seems inappropriate in this brief and selective review to do more than mention the obviously important contribution first of protein sequencing, later of RNA and cDNA sequencing to understanding of the influenza virus genome and its products—notably the hemagglutinin. This basic information has permitted not only the intracistronic mapping of hemagglutinin mutants by analysis of tryptic peptides (Laver *et al.*, 1980) but a beginning identification of antigenic sites (Laver *et al.*, 1980; Lubeck and Gerhard, 1981; Sleigh *et al.*, 1981; Wiley *et al.*, 1981) in association with the selective use of monoclonal antibodies (see Chapters 2, 4, and 5).

III. Viral Genetics and the Understanding of Viral Virulence and Pathogenicity

With segmented genome viruses, the selective reshuffling of genes provides an elegant system for identifying those genes contributing to that complex phenomenon which we define as viral virulence, the capacity of a virus to injure the host in some readily perceived manner. As I have pointed out, the earliest markers available for the study of influenza virus genetics involved virulence, and were for the most part polygenic. Therefore, it is not surprising that virulence remains a somewhat evanescent endpoint, the definition of which is still difficult even with the more precise methods of viral gene identification now at hand. The problems in the genetic dissection of virulence by "cut and paste" reassortment genetics include 1) the possibility of mutation occurring in reassorted genes during gene reassortment or during isolation of the reassortant virus (Erickson and Kilbourne, 1980; Campbell *et al.*, 1982), 2) the acquisition of new viral phenotypes as genes are reassembled in new contexts or "gene constellations" (Laver and Kilbourne, 1966; Rott *et al.*, 1976), and 3) the need for precise definition and assessment of virulence endpoints, especially in complex (*i.e.*, intact) hosts. Nevertheless, important contributions have been made even as the limitations of the analytical systems have been recognized (see Chapter 9).

In discussing genetic reassortment as a means of generating an attenuated live virus vaccine, it was predicted that recombinants of wild type (virulent) and laboratory (avirulent) strains would, on a statistical basis, be intermediate in virulence between the two parents, assuming that virulence was dependent on several genes (Kilbourne, 1969) and that genetic reassortment was random. For the most part, this prediction has proved correct (Beare and Hall, 1971; Webster *et al.*, 1971; Mayer *et al.*, 1973; Rott *et al.*, 1976; Scholtissek *et al.*, 1977; Florent *et al.*, 1977; Rott *et al.*, 1978; Bean and Webster, 1978), but experimental genetic reassortment has occasionally produced highly neurovirulent recombinants derived from nonneurovirulent parent viruses (Scholtissek *et al.*, 1979) and, conversely, nonpathogenic recombinants from highly pathogenic strains (Rott *et al.*, 1979).

A reductionist approach has led to the identification of single genes which determine virulence in limited, explicitly defined laboratory systems. But virulence here often relates to plaque formation in monolayer cultures (Almond, 1977; Israel, 1980). Furthermore, because of a requirement for hemagglutinin cleavage before infection can incur, hemagglutinin mutants may be polar in the sense that they determine as a single gene this critical initial event. In turn, where the nature of the neuraminidase determines whether cleavage will occur, this gene can be critical for the initiation of viral replication leading to pathogenic effects (Schulman and Palese, 1977). Point mutations carried by two temperature sensitive mutants are attenuating for virus infecting man (Massicot *et al.*, 1980) and mutation in a single swine influenza virus (hemagglutinin) gene pleiotropically influences antigenicity, replication capacity in chick embryos and MDCK cells and infectivity for the virus' natural host (Kilbourne, 1978; Kilbourne, McGregor, and Easterday, 1979).

IV. Influenza Virus Genetics and the Epidemiology and Evolution of Influenza Viruses (Molecular Epidemiology)

The identification of the hemagglutinin and neuraminidase genes as separate and reassortable entities led to the prompt recognition in the pandemic of 1968 that the new Hong Kong virus contained a novel H3 hemagglutinin and a neuraminidase (N2) left over from the preceding decade (Schulman and Kilbourne, 1969). Consideration of this experience in the context of serologic evidence for the recycling of influenza viral antigens prompted introduction of the term "molecular epidemiology" (Kilbourne, 1974, 1979), based on evidence for reassortment, in nature, of genes for the surface antigens of influenza viruses. In the past decade, as all influenza viral genes have become available for scrutiny by molecular analysis, a true science of molecular epidemiology has developed for many animal viruses, but most notably for the herpes (Buchman *et al.*, 1978) and influenza viruses (Palese *et al.*, 1981). RNA-RNA hybridization has now shown the probable derivation not only of the neuraminidase but of all but the hemagglutinin gene of H3N2 virus from the antecedent human H2N2 virus (Scholtissek *et al.*, 1978), the hemagglutinin gene having apparently been acquired from an animal source by reassortment (Scholtissek *et al.*, 1978; Laver *et al.*, 1973; Fang *et al.*, 1981; Ward and Dopheide, 1981).

There is mounting evidence that genetic reassortment occurs not only among animal influenza viruses (reviewed by Webster and Laver, 1975; Desselberger *et al.*, 1978) but also in man (Young, Desselberger, and Palese, 1979; Bean *et al.*, 1980). Of equal interest is evidence from oligonucleotide maps that the H1N1 human virus which appeared abruptly in 1977 is closely related to virus which circulated in the early 1950's (Nakajima, Desselberger, and Palese, 1978). Studies of this newly emergent virus as it spread throughout China, its country of apparent origin, suggested sequential mutations in the virus with later strains originating by divergent evolution from a common ancestor (Young, Desselberger, and Palese, 1979).

A. Genetic Reassortment in Nature and Its Contribution to the Evolution of New Viruses

It has been tempting to relate the puzzling epidemiology of influenza to the striking capacity of the virus to undergo genetic reassortment in the laboratory. Early speculation that the novel major variants associated with pandemics might be derived by recombination of human and animal viruses (Rasmussen, 1964; Kilbourne, 1968, 1974; Webster *et al.*, 1971) has received increasing support through the years. Webster and his colleagues provided experimental evidence that genetic reassortment of swine, avian and human viruses could occur under conditions of contact exposure simulating "natural" conditions (Webster *et al.*, 1973). Furthermore, the variegated mixture of hemagglutinin and neuraminidase antigens found in nature, especially among avian viruses, together with evidence that dual infection is relatively frequent (Hinshaw *et al.*, 1980) demonstrated that conditions exist-

ing in nature are suitable for genetic reassortment. Oligonucleotide analysis of two avian field isolates (Hav6N2 and Hav6Nav4) demonstrated that they possessed nearly identical M and HA genes but that other genes were markedly dissimilar. It was concluded that these viruses must be related by a recombinational event (Desselberger *et al.*, 1978).

Until 1976 the coincident circulation of more than one subtype of influenza A virus in man was not observed. From the intense outbreak of influenza at Fort Dix, New Jersey in that year, both H1N1 (swine) and H3N2 viruses were isolated, but evidence of reassortment between these strains was not found by viral genotyping by RNA PAGE.

However, the protracted co-circulation of "Russian" H1N1 and H3N2 viruses in man in subsequent years has resulted in demonstrable double infection (Kendal *et al.*, 1979) and in the emergence of reassortant viruses with significant epidemic potential (Young and Palese, 1979; Bean *et al.*, 1980).

B. The Genetics of Minor Variation

The minor variation of influenza viruses in interpandemic periods results from sequential point mutations in the viral RNA (reviewed by Kilbourne, 1975; Young, Desselberger, and Palese, 1979; Both *et al.*, 1980; Laver *et al.*, 1980) which lead to antigenic changes in both hemagglutinin and neuraminidase (Schulman and Kilbourne, 1969) and probably in some changes in all viral proteins (Palese and Young, 1982). More than antigenic variation seems to be required for survival of the virus (Kilbourne, 1980) and changes in virulence for man, although difficult to assess, undoubtedly occur (Semple, 1951).

Within a local epidemic, oligonucleotide mapping of total genome RNA disclosed two to four nucleotide changes in all strains examined over a period of two months. When compared to a strain recovered in the same city one year later, 25 changes were noted from the 1976 reference strains (Palese *et al.*, 1981). Over a decade of interpandemic prevalence, oligonucleotide mapping of the HA gene alone of five strains, revealed nucleotide changes of 0.96 to 3.4 percent, increasing with time (Palese *et al.*, 1981). Studying cDNA copies of virion RNA of the H3 hemagglutinin, Both and Sleigh (1981) found 36 amino acid changes within four variable regions in little more than a decade after the appearance of this subtype.

C. Influenza A, B, and C Viruses

Originally classified together as influenza or orthomyxoviruses on the basis of the similar clinical syndromes which they induce, it is now clear that the similarities among these viruses far exceed their differences (see Review by Palese and Young, 1982). All are segmented genome enveloped viruses which can undergo genetic reassortment (Perry *et al.*, 1954; Tobita and Kilbourne, 1974; Racaniello and Palese, 1979). Reassortment between any two types has not been demonstrated. However, a high degree of homology is observed when terminal RNA sequences of influenza

A, B, and C viruses are compared, and inverted partial complementarity of sequences of the 5′ and 3′ ends of all genes is found (Desselberger *et al.*, 1980). In accord with epidemiological observations, cDNA-RNA hybridization analysis of RNA's of the three viral types reveals decrease in variation from A to C influenza viruses (Palese *et al.*, 1981).

In evolutionary terms it may be that influenza C viruses are the more obligate human pathogens of the three types. Only recently has infection of animals in nature been described (Kuo *et al.*, 1982) and transmission or establishment of the virus in animals has not been observed. Large epidemics of influenza C in man do not occur, perhaps related to the relative avirulence of this virus. Because, therefore infection in childhood is sporadic, not involving all children simultaneously, influenza C viruses may survive through serial propagation in susceptibles without requiring the frequent selection of new antigenic variants needed to surmount the ubiquitous antibody barriers that confront influenza A viruses.

V. Practical Applications of Influenza Viral Genetics

The first practical application of virus genetics to the design of vaccines followed the discovery (Kilbourne and Murphy, 1960) that the capacity of an established laboratory strain to replicate efficiently in the standard chick embryo host could be combined by genetic reassortment with the desired antigenic characteristics of a newly isolated poorly growing virus in a high yielding reassortant. Neither regulatory authorities nor vaccine manufacturers were quick to accept this new approach to the rapid production of vaccine strains produced by genetic manipulation, so that not until 1971 was X-31, a recombinant of PR8 and the 1968 H3N2 Aichi strain, extensively tested in volunteers (Couch *et al.*, 1971) and field trials (Leibovitz *et al.*, 1971), then later used in commercial vaccine production (Baez, Palese, and Kilbourne, 1980). Subsequent studies confirmed serologic evidence that this first genetically engineered vaccine virus contained H3N2 glycoprotein genes but all other genes from its PR8 parent (Baez, Palese, and Kilbourne, 1980), and that all other high yielding vaccine strains genotyped by RNA electrophoresis derived at least the M gene from PR8.

Examining recombinants of PR8 and Hong Kong (HK) virus, Schulman and Palese (1978) and found M and NP genes from PR8 to be associated with high yields. Virtually all influenza A vaccine viruses are now routinely produced by genetic reassortment of new viruses with PR8 or X-31 viruses.

Genetic reassortment has also been used to transfer defined temperature sensitive *ts* or *ca* lesions to live virus vaccine strains (reviewed by Stuart-Harris, 1980). In this case, of course, prior genetic manipulation had led to isolation and characterization of appropriate *ts* or *ca* mutants.

Reassortant high yielding viruses bearing the desired HA or NA genes have been valuable in biochemical studies of influenza viruses, as is notably the case with X-31, which has been used as a source of abundant hemagglutinin and RNA for protein and RNA sequencing studies and most recently for definition of the ultrastructure of the hemagglutinin molecule (Wilson *et al.*, 1981).

Antigenic hybrids have, from the first, been useful in the production of what are effectively monospecific antisera if the HA or NA of the immunizing virus is matched only to one or the other antigens in the test virus (Kilbourne *et al.*, 1968). Such hybrids may also permit physical separation of antigens not possible when they are incorporated in the parental virus (Laver and Kilbourne, 1966).

Antigenic hybrids also have been used in the induction of neuraminidase-specific infection-permissive immunity in man as a new approach to influenza vaccination (Couch *et al.*, 1974).

VI. The Special Genetics of Divided Genome Viruses in Relation to Problems in Influenza

Five different groups of RNA viruses of animals possess divided or segmented genomes, including such dissimilar prototypes as Drosophila X, bluetongue, Bunyamwera and infective pancreatic necrosis viruses (reviewed by Kilbourne, 1981). It is possible that such viruses have an intrinsic evolutionary and adaptive advantage over viruses competing within the same ecosystem. These viruses are not limited by mutational frequency for their variation but have available an extended gene pool carried by related viruses with which they can carry out genetic reassortment. In the case of influenza A viruses, this genetic reservoir exists in certain domestic animals and wild birds. From this extended pool, gene combinations optimal for virulence, transmissibility and antigenic fit can be continuously sampled to the extent that interspecific transfer of viruses may occasionally occur (Kilbourne, 1973, 1979).

Genetic reassortment provides a mechanism analogous to sexual reproduction for the testing of new mutations in different contexts. It has been pointed out that the greater the variation in an asexual population, the less likely such a mutation is to be part of the best genotype. Sex (or reassortment) allows a mutation to be tested in many genotypes and to be selected according to its average contribution to fitness rather than according to the fitness of the genotype whence it originates (Smith, 1978). Thus, a primary result of genetic reassortment is the simultaneous fixation of different favorable mutations in a new virus. Smith (1978) has also pointed out that when selection is intense, sexual populations (here, analogously, reassorting heterogeneous viruses) evolve much more rapidly than asexual (read non-reassorting) ones.

Genetic reassortment, of course, provides no obvious answer for the progressive minor antigenic variation of influenza viruses in nature. However, it could provide a *de facto* increase in mutational frequency by the salvage of potentially lost mutations occurring in noninfective particles by their incorporation through endogenous recombination and recruitment in multicycle infection into virions. Thus an extended mutational pool would be created which could enhance the opportunity for selection of advantageous mutations—antigenic or otherwise (Kilbourne, 1973, 1981).

A new dimension to evolution is the cold storage genetic reservoirs represented by laboratory preservation of supposedly extinct viruses that have apparently

disappeared in nature. No latency mechanism allowing natural persistence is known for influenza viruses, but isolated genes conceivably might be preserved following infection for later rescue or recapture in reassortant viruses. This theory does not readily explain the reemergence of the 1977 H1N1 virus, the genome of which seemed almost identical with an early 1950 strain (Nakajima, Desselberger, and Palese, 1978). In any case, to the extent that strains in the laboratory are potentially part of the total ecologic reservoir of the virus, the gene pool of influenza viruses is extended to another time frame, including all viruses (and genes) isolated and preserved since 1901 (Kilbourne, 1981), the date fowl plague virus was first isolated. However, the constraints and barriers to the laboratory as a source of infection are many (Kilbourne, 1974).

VII. Unsolved Problems in Influenza and Genetic Approaches to Their Solution

While influenza remains as an unsolved problem at the end of the 20th Century because of the genetic resources of its virus, it is possible that these very resources may be used for its control and eventual eradication. The utilization of genetic manipulation for production of conventional and experimental vaccines has already been reviewed. Beyond this, the sequencing of influenza virus genes promises to yield valuable information on the origin and interrelationships of viruses. The independent synthesis of isolated gene products will allow study of their biological functions and understanding of their contributions to virulence and toxicity. Thus, modified hemagglutinins might function immunogenically but be devoid of any toxicity. Immunogenicity might be enhanced by steric or other alterations in molecular structure which could be induced by whole or partial gene synthesis.

Ultimately, it is not beyond reason that the virulent viruses in nature might be replaced by highly transmissible but avirulent strains. But, of course, simultaneous limitation of their interaction with other influenza viruses in animal reservoirs would be essential.

References

Air, G. M.: Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus. Proc. Natl. Acad. Sci. *78*, 7639–7643 (1981).

Almond, J. W. (1977): A single gene determines the host range of influenza virus. Nature *270*, 617–618 (1977).

Baez, M., Palese, P., Kilbourne, E. D.: Gene composition of high-yielding influenza vaccine strains obtained by recombination. J. Infect. Dis. *141*, 362–365 (1980).

Baron, S., Jensen, K. E.: Evidence for genetic interaction between noninfections and infections influenza A viruses. J. Exp. Med. *102*, 677–697 (1955).

Barry, R. D.: Study of infectivity and hemagglutination of influenza virus in deembryonated eggs. J. Immunol. *65*, 571–583 (1950 a).

Barry, R. D.: The multiplication of influenza virus. II. Multiplicity reactivation of ultraviolet irradiated virus. Virology *14*, 398–405 (1961 b).

Bean, W.J.: Recombination of human influenza A viruses in nature. Nature *284,* 638–640 (1980).
Bean, W.J., Webster, R.G.: Phenotypic properties associated with influenza genome segments. In: Negative Strand Viruses and the Host Cell (Mahy, B.W.J., Barry, R.D., eds.), 685–692. New York: Academic Press 1978.
Beare, A.S., Hall, T.S.: Recombinant influenza-A viruses as live vaccines for man. Lancet *2,* 1271–1273 (1971).
Both, G.W., Sleigh, M.J.: Conservation and variation in the hemagglutinins of Hong Kong subtype influenza viruses during antigen drift. J. Virol. *39,* 663–672 (1981).
Both, G.W., Sleigh, M.J., Bender, V.J., Moss, B.A.: A comparison of antigenic variation in Hong Kong influenza virus haemagglutinins at the nuclei acid level. In: Structure and Variation in Influenza Virus (Laver, W.G., Air, G.M., eds.), 81–89. New York: Elsevier/North-Holland 1980.
Briody, B.A., Cassell, W.A., Medill, M.A.: Adaptation of influenza virus to mice. III. Development of resistance to inhibitor. J. Immunol. *74,* 41–45 (1955).
Buchmann, T.G., Roizman, B., Adams, G., Stover, B.H.: Restriction endonuclease fingerprinting of herpes simplex virus DNA: A novel epidemiological tool applied to a nosocomial outbreak. J. Infect. Dis. *138,* 488–498 (1978).
Burge, B.W., Pfefferkorn, E.R.: Isolation and characterization of conditional lethal mutants of sindbis virus. Virology *30,* 204–213 (1966).
Burnet, F.M.: Influenza virus infections of the chick embryo by the amniotic route. I. General character of the infections. Austral. J. Exp. Biol. Med. Sci. *18,* 353–360 (1940).
Burnet, F.M.: Structure of influenza virus. Science *123,* 1101–1104 (1956).
Burnet, F.M., Bull, D.R.: Changes in influenza virus associated with adaptation to passage in chick embryos. Austral. J. Exp. Biol. Med. Sci. *21,* 55–69 (1943).
Burnet, F.M., Lind, P.E.: Recombination of characters between two influenza virus strains. Austral. J. Sci. *12,* 109–110 (1949).
Burnet, F.M., Lind, P.E.: Reactivation of heat inactivated influenza virus by recombination. Austral. J. Exp. Biol. Med. Sci. *32,* 133–143 (1954 a).
Burnet, F.M., Lind, P.E.: An analysis of the adaptation of an influenza virus to produce lesions in the mouse lung. Austral. J. Exp. Biol. Med. Sci. *32,* 711–720 (1954 b).
Campbell, D., Sweet, C., Hay, A.J., Douglas, A., Skehel, J.J., Mason, T.J., Smith, H.: Genetic composition and virulence of influenza virus: Differences in facets of virulence in ferrets between two pairs of recombinants with RNA segments of the same parental origin. J. Gen. Virol. *58,* 387–398 (1982).
Couch, R.B., Douglas, R.G., jr., Fedson, D.S., Kasel, J.A.: Correlated studies of a recombinant influenza-virus vaccine. III. Protection against experimental influenza in man. J. Infect. Dis. *124,* 473–480 (1971).
Couch, R.B., Kasel, J.A., Gerin, J.L., Schulman, J.L., Kilbourne, E.D.: Induction of partial immunity to influenza by a neuraminidase-specific vaccine. J. Infect. Dis. *129,* 411–420 (1974).
Desselberger, U., Racaniello, V.R., Zazra, J., Palese, P.: The 3′ and 5′ terminal sequences of influenza A, B, and C virus genes are highly conserved and show partial inverted complementarity. Gene *8,* 315–328 (1980).
Desselberger, U., Zamecnik, P., Palese, P.: 3′-Terminal sequences of hemagglutinin and neuraminidase genes of different influenza A viruses. In: Proceedings of the International Workshop on Structure and Variation in Influenza Virus, Thredbo, Australia (Laver, W.G., Air, G.M., eds.), 169–179. New York: Elsevier/North-Holland 1980.
Duesberg, P.H.: The RNA's of influenza virus. Proc. Natl. Acad. Sci. *59,* 930–937 (1968).
Erickson, A.H., Kilbourne, E.D.: Mutation in the hemagglutinin of A/N-WS/33 influenza virus recombinants influencing sensitivity to trypsin and antigenic reactivity. Virology *107,* 320–330 (1980).
Fang, R., Jou, W.M., Huylebroeck, D., Devos, R., Fiers, W.: Complete structure of A/duck/Ukraine/63 influenza hemagglutinin gene: Animal virus as progenitor of human H3 Hong Kong 1968 influenza hemagglutinin. Cell *2,* 315–323 (1981).
Fenner, F.: Conditional lethal mutants in the study of the genetics of animal viruses. In: Perspectives in Virology IV (Pollard, M., ed.), 34–46. New York: Harper & Row 1965.
Florent, G., Lobmann, M., Beare, A.S., Zygraich, N.: RNAs of influenza virus recombinants derived from parents of known virulence for man. Arch. Virol. *54,* 19–28 (1977).

Francis, T., jr., Moore, A. E.: A study of the neurotropic tendency in strains of the virus of epidemic influenza. J. Exp. Med. *72,* 717–745 (1940).

Gottlieb, T., Hirst, G. K.: The experimental production of combination forms of virus. III. The formation of doubly antigenic particles from influenza A and B virus and a study of the ability of individual particles of X virus to yield separate strains. J. Exp. Med. *99,* 307–320 (1954).

Henle, W., Liu, O. C.: Studies on host-virus interactions in the chick embryo-influenza virus system. VI. Evidence for multiplicity reactivation of inactivated virus system. J. Exp. Med. *94,* 305–322 (1951).

Hinshaw, V. S., Bean, W. J., Webster, R. B., Siram, G.: Genetic reassortment of influenza A viruses in the intestinal tract of ducks. Virology *102,* 412–419 (1980).

Hirst, G. K.: Agglutination of red cells by allantoic fluid of chick embryos infected with influenza virus. Science *94,* 22–23 (1941).

Hirst, G. K.: Genetic recombination with newcastle disease virus, polioviruses, and influenza. Cold Spring Harbor Symp. on Quant. biol. *27,* 303–309 (1962).

Hirst, G. K.: Mechanism of influenza recombination. I. Factors influencing recombination rates between temperature-sensitive mutants of strain WSN and the classification of mutants into complementation-recombination groups. Virology *55,* 81–93 (1973).

Hirst, G. K., Gotlieb, T.: The experimental production of combination forms of virus. II. A study of serial passage in the allantoic sac of agents that combine the antigens of two distinct influenza A strains. J. Exp. Med. *98,* 53–70 (1953).

Holland, J., Spindler, K., Horodysky, F., Grabau, E., Nichol, S., Vande Pol, S.: Rapid evolution of RNA genomes. Science *215,* 1577–1585 (1982).

Horsfall, F. L.: Reproduction of influenza viruses. Quantitative investigations with particle enumeration procedures on the dynamics of influenza A and B virus reproduction. J. Exp. Med. *102,* 441–473 (1955).

Hoyle, L.: The influenza viruses. In: Gard, S., Hallauer, C., Meyer, K. F. (eds.), Virology Monographs, Vol. 4. Wien-New York: Springer 1968.

Isaacs, A., Edney, M.: Variation in laboratory stocks of influenza viruses: Genetic aspects of the variations. Brit. J. Exp. Path. *31,* 209–216 (1950).

Israel, A.: Genotypic and phenotypic characterization of a mammalian cell-adapted mutant of fowl plague virus (FPV). J. Gen. Virol. *51,* 33–44 (1980).

Jahiel, R. I., Kilbourne, E. D.: Reduction in plaque size and reduction in plaque number as differing indices of influenza virus-antibody reactions. J. Bacteriol. *92,* 1521–1534 (1966).

Jou, W. M., Verhoeyen, M., Debos, R., Saman, E., Fang, R., Huylebroeck, D., Fiers, W., Threlfall, G., Barber, C., Carey, N., Emtage, S.: Complete structure of the hemagglutinin gene from the human influenza A/Victoria/3/75 (H3N2) strain as determined from cloned DNA. Cell *19,* 683–696 (1980).

Kendal, A. P., Lee, D. T., Parish, H. S., Raines, D. Noble, G. R., Dowdle, W. R.: Laboratory-based surveillance of influenza virus in the United States during the winter of 1977–1978. II. Isolation of a mixture of A/Victoria- and A/USSR-like viruses from a single person during an epidemic in Wyoming, U.S.A., January 1978. Amer. J. Epidemiol. *110,* 462–468 (1979).

Kilbourne, E. D.: Reactivation of non-infective virus in a cortisone injected host. J. Exp. Med. *101,* 437–450 (1955).

Kilbourne, E. D.: The influence of cortisone on experimental viral infection. IV. Negation of interference as the mechanism by which cortisone induces increased virus yields. J. Exp. Med. *106,* 863–881 (1957).

Kilbourne, E. D.: Influenza virus genetics. Progr. Med. Virol. *5,* 79–126 (1963).

Kilbourne, E. D.: Recombination of influenza A viruses of human and animal origin. Science *160,* 74–76 (1968).

Kilbourne, E. D.: The molecular epidemiology of influenza. J. Infect. Dis. *127,* 478–487 (1973).

Kilbourne, E. D.: An explanation of the interpandemic antigenic mutability of influenza viruses. J. Infect. Dis. *128,* 668–670 (1973).

Kilbourne, E. D.: Virus research (letter). Science *184,* 410–411 (1974).

Kilbourne, E. D.: Epidemiology of influenza. In: The Influenza Viruses and Influenza (Kilbourne, E. D., ed.), 483–538. New York: Academic Press 1975.

Kilbourne, E. D.: Genetic dimorphism in influenza viruses: Characterization of stably associated hemagglutinin mutants differing in antigenicity and biological properties. Proc. Natl. Acad. Sci. *75,* 6258–6262 (1978).

Kilbourne, E. D.: Molecular epidemiology-influenza as archetype. The Harvey Lectures *73,* 225–258 (1979).

Kilbourne, E. D.: Influenza: Viral determinations of the pathogenicity and epidemicity of an invariant disease of variable occurrence. Phil. Trans. R. Soc. (Lond.) *B 288,* 291–297 (1980).

Kilbourne, E. D.: Recent contributions of molecular biology to the clinical virology of myxoviruses. Yale J. Biol. Med. *53,* 41–45 (1980).

Kilbourne, E. D.: Segmented genome viruses and the evolutionary potential of asymmetrical sex. Perspect. Biol. Med. *25,* 66–77 (1981).

Kilbourne, E. D., Choppin, P. W., Schulze, I. T., Scholtissek, C., Bucher, D. L.: Influenza virus polypeptides and antigens-summary of influenza workshop I. J. Infect. Dis. *125,* 447–455 (1972).

Kilbourne, E. D., Laver, W. G., Schulman, J. L., Webster, R. G.: Antiviral activity of antiserum specific for an influenza virus neuraminidase. J. Virol. *2,* 281–288 (1968).

Kilbourne, E. D., Lief, F. S., Schulman, J. L., Jahiel, R. I., Laver, W. G.: Antigenic hybrids of influenza viruses and their implications. Perspect. Virology *5,* 87–106 (1967).

Kilbourne, E. D., McGregor, S., Easterday, B. C.: Hemagglutinin mutants of swine influenza virus differing in replication characteristics in their natural host. Infect. Immun. *26,* 197–201 (1979).

Kilbourne, E. D., Murphy, J. S.: Genetic studies of influenza viruses. I. Viral morphology and growth capacity as exchangeable genetic traits. Rapid *in ovo* adaptation of early passage asian strain isolates by combination with PR8. J. Exp. Med. *111,* 387–406 (1960).

Kilbourne, E. D., Schulman, J. L.: The induction of broadened (multitypic) immunity with doubly antigenic influenza virus recombinants. Trans. Assoc. Am. Physicians *78,* 323–333 (1965).

Kilbourne, E. D., Schulman, J. L., Schild, G. C., Schloer, G., Swanson, J., Bucher, D.: Correlated studies of a recombinant influenza-virus vaccine. I. Derivation and characterization of virus and vaccine. J. Infect. Dis. *124,* 449–462 (1971).

Klenk, H.-D., Rott, R., Orlich, M., Blodorn, J.: Activation of influenza A viruses by Trypsin treatment. Virology *68,* 426–439 (1975).

Koprowski, H., Gerhard, W.: In: Animal Virus Genetics (Fields, B., Jaenisch, R., Fox, C., eds.), 591. New York: Academic Press 1980.

Kuo, Y. C., Fengen, J., Min, W., Ping, W., Chu, C. M.: Influenza C virus isolated from pigs in China. Kexue Tongbao *27,* 1118–1121 (1982).

Laver, W. G.: The structure of influenza viruses. 3. Disruption of the virus particle and separation of neuraminidase activity. Virology *20,* 251–262 (1963).

Laver, W. G.: Structural studies on the protein subunits from three strains of influenza virus. J. Mol. Biol. *9,* 109–124 (1964).

Laver, W. G., Air, G. M., Dopheide, T. A., Ward, C. W.: Amino acid sequence changes in the hemagglutinin of A/Hong Kong (H3N2) influenza virus during the period 1968–77. Nature *283,* 454–457 (1980).

Laver, W. G., Kilbourne, E. D.: Identification in a recombinant influenza virus of structural proteins derived from both parents. Virology *30,* 493–501 (1966).

Laver, W. G., Webster, R. G.: Selection of antigenic mutants of influenza viruses. Isolation and peptide mapping of their hemagglutinating proteins. Virology *34,* 193–202 (1968).

Laver, W. G., Webster, R. G.: Studies on the origin of pandemic influenza. III. Evidence implicating duck and equine influenza viruses as possible progenitors of the Hong Kong strain of human influenza. Virology *51,* 383–391 (1973).

Lazarowitz, S. G., Choppin, P. W.: Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. Virology *68,* 440–454 (1975).

Leibovitz, A., Coultrip, R. L., Kilbourne, E. D., Legters, L. J., Smith, C. D., Chin, J., Schulman, J. L.: Correlated studies of a recombinant influenza-virus vaccine. IV. Prodection against naturally occurring influenza in military trainees. J. Infect. Dis. *124,* 481–487 (1971).

Lubeck, M. D., Gerhard, W.: Topological mapping of antigenic sites on the influenza A/PR/8/34 virus hemagglutinin using monoclonal antibodies. Virology *113,* 64–72 (1981).

Massicot, J. G., Murphy, B. R., Van Wyke, K., Huan, K-Y., Chanock, R. M.: ts P1 and P3 genes are

responsible for satisfactory level of attenuation of ts-1A2 recombinants bearing H1N1 or H3N2 surface antigenes of influenza A virus. Virology *106,* 187–190 (1980).

Massicot, J. G., Van Wyke, K., Chanock, R. B., Murphy, B. R.: Evidence for intrasegmental complementation between two influenza A viruses having ts mutations on their P1 genes. Virology *117,* 496–500 (1982).

Mayer, V., Schulman, J. L., Kilbourne, E. D.: Nonlinkage of neurovirulence exclusively to viral hemagglutinin or neuraminidase in genetic recombinants of A/NWS (H0N1) influenza virus. J. Virol. *11,* 272–278 (1973).

McClelland, A. J., Hare, J. D.: Adsorption of influenza virus by red cells and a new *in vitro* method of measuring antibodies for influenza virus. Canad. Publ. Health J. *32,* 530–538 (1941).

Murphy, B. R., Tolpin, M. D., Massicot, J. G., Kim, H. Y., Parrott, R. H., Chanock, R. M.: Escape of a highly defective influenza A virus mutant from its temperature sensitive phenotype by extragenic suppression and other types of mutation. Annals New York Acad. Sci. *354,* 172–182 (1980).

Nakajima, K., Desselberger, U., Palese, P.: Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. Nature *274,* 334–339 (1978).

Nakajima, K., Sugiura, A.: Three-factor cross of influenza virus. Virology *81,* 486–489 (1977).

Nottay, B. K., Kew, O. M., Hatch, M. H., Heyward, J. T., Obijeski, J. F.: Molecular variation of type 1 vaccine-related and wild polioviruses during replication in humans. Virology *108,* 405–423 (1981).

Palese, P., Brand, C., Young, J. F., Baez, M., Six, H. R., Kasel, J. A.: Molecular epidemiology of influenza viruses. Perspect. Virology *11,* 115–127 (1981).

Palese, P., Elliott, R. M., Baez, M., Zazra, J. J., Young, J. F.: Genome diversity among influenza A, B and C viruses and genetic structure of RNA 7 and RNA 8 of influenza A viruses. In: Genetic Variation Among Influenza Viruses (Nayak, D., ed.), 127–140. New York: Academic Press 1981.

Palese, P., Schulman, J. L.: Differences in RNA pattern of influenza viruses. J. Virol. *17,* 876–884 (1976).

Palese, P., Young, J. F.: Variation of influenza A, B and C viruses. Science *215,* 1468–1474 (1982).

Perry, B. T., Van Den Ende, M., Burnet, F. M.: Recombination with two influenza B strains in the deembryonated egg. Austral. J. Exp. Biol. *32,* 469–478 (1954).

Pons, M. W., Hirst, G. K.: Polyacrylamide gel electrophoresis of influenza virus RNA. Virology *34,* 385–388 (1968).

Racaniello, V. R., Palese, P.: Isolation of influenza C virus recombinants. J. Virol. *32,* 1006–1014 (1979).

Rasmussen, jr., A. F.: Avian myxoviruses and man. In: Newcastle Disease Virus: An Evolving Pathogen (Hanson, R. P., ed.), 313–325. Madison: University Wisconsin Press 1964.

Ritchey, M., Palese, P.: *In vitro* translation of influenza virus messenger RNAs. Virology *72,* 410 (1976).

Ritchey, M. B., Palese, P., Schulman, J. L.: Differences in protein patterns of influenza A viruses. Virology *76,* 122–128 (1977).

Rohde, W., Scholtissek, C.: On the origin of the gene coding for an influenza A virus nucleocapsid protein. Arch. Virol. *64,* 213–223 (1980).

Rott, R., Orlich, M., Scholtissek, C.: Attenuation of pathogenicity of fowl plague virus by recombination with other influenza A viruses nonpathogenic for fowl: nonexclusive dependence of pathogenicity on hemagglutinin and neuraminidase of the virus. J. Virol. *19,* 54–60 (1976).

Rott, R., Orlich, M., Scholtissek, C.: Correlation of pathogenicity and gene constellation of influenza A viruses. III. Non-pathogenic recombinants derived from highly pathogenic parent strains. J. Gen. Virol. *44,* 471–477 (1979).

Rott, R., Scholtissek, C., Klenk, H–D., Orlich, M.: In: Negative Strand Viruses and the Host Cell (Many, B. W. J., Barry, R. D., eds.), 652–662. New York: Academic Press 1978.

Scholtissek, C., Rohde, W., Harms, E., Rott, R., Orlich, M., Boschek, C. B.: A possible partial heterozygote of an influenza A virus. Virology *89,* 506–516 (1978).

Scholtissek, C., Rohde, W., Von Hoyningen, V., Rott, R.: On the origin of the human influenza virus subtypes H2N2 and H3N2. Virology *87,* 13–20 (1978).

Scholtissek, C., Rott, R., Orlich, M., Harms, E., Rohde, W.: Correlation of pathogenicity and gene constellation of an influenza A virus (fowl plaque). I. Exchange of a single gene. Virology *82,* 74–80 (1977).

Scholtissek, C., Vallbracht, A., Flehmig, B., Rott, R.: Correlation of pathogenicity and gene constellation of influenza A viruses. II. Highly neurovirulent recombinants derived from non-neurovirulent or weakly neurovirulent parent virus strains. Virology *95,* 492–500 (1979).

Schulman, J. L., Kilbourne, E. D.: Experimental influenza virus infection of mice. Proc. First Intern. Symp. Aerobiol., 141–146. University of California Press 1963.

Schulman, J. L., Kilbourne, E. D.: Independent variation in nature of hemagglutinin and neuraminidase antigens of influenza virus: Distinctiveness of hemagglutinin antigen of Hong Kong/68 virus. Proc. Natl. Acad. Sci. *63,* 326–333 (1969).

Schulman, J. L., Palese, P.: Selection and identification of influenza virus recombinats of defined genetic composition. J. Virol. *20,* 248–254 (1976).

Schulman, J. L., Palese, P.: Virulence factors of influenza A viruses. WSN virus neuraminidase required for productive infection in MDCK cells. J. Virol. *24,* 170–176 (1977).

Schulman, J. L., Palese, P.: Biological properties of recombinants of influenza A/Hong Kong and A/PR8 viruses: Effects of genes for matrix protein and nucleoprotein on virus yield in embryonated eggs. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 663–674. New York: Academic Press 1978.

Semple, A. B.: Epidemiology of the influenza epidemic in Liverpool in 1950/51. Proc. Roy. Soc. Med. *44,* 794–796 (1951).

Simpson, R. W., Hirst, G. K.: Genetic recombination among influenza viruses. I. Cross reactivation of plaque-forming capacity as a method for selecting recombinants from the progeny of crosses between influenza A strains. Virology *15,* 436–451 (1961).

Simpson, R. W., Hirst, G. K.: Temperature-sensitive mutants of influenza A virus: Isolation of mutants and preliminary observations on genetic recombination and complementation. Virology *35,* 41–49 (1968).

Sleigh, M. J., Both, G. W., Underwood, P. A., Bender, V. J.: Antigenic drift in the hemagglutinin of the Hong Kong influenza subtype: Correlation of amino acid changes with alterations in viral antigenicity. J. Virol. *37,* 845–853 (1981).

Smith, J. M.: The Evolution of Sex. New York: Cambridge Univ. Press. 1978.

Smith, W.: Cultivation of the virus of influenza. Brit. J. Exp. Path. *16,* 508–512 (1935).

Spring, S. B., Nusinoff, S. R., Tierney, E. L., Richman, D. D., Murphy, B. R., Chanock, R. M.: Temperature-sensitive mutants of influenza. VIII. Genetic and biological characterizations of ts mutants of influenza virus A (H3N2) and their assignment to complementation groups. Virology *66,* 542–550 (1975).

Staiger, H. R.: Plaque-type recombination in fowl plague virus. Virology *22,* 419–422 (1964).

Stuart-Harris, C.: A neurotropic strain of human influenza virus. Lancet *i,* 497–499 (1939).

Stuart-Harris, C.: The present status of live influenza virus vaccine. J. Infect. Dis. *142,* 784–793 (1980).

Sugiura, A.: Influenza virus genetics. In: The Influenza Viruses and Influenza (Kilbourne, E. D., ed.), 171–213. New York: Academic Press 1975.

Sugiura, A., Kilbourne, E. D.: Genetic studies of influenza viruses. II. Plaque formation by influenza viruses in a clone of a variant human heteroploid cell line. Virology *26,* 478–488 (1965).

Sugiura, A., Kilbourne, E. D.: Genetic studies of influenza viruses. III. Production of plaque-type recombinants with A_0 and A_1 strains. Virology *29,* 84–91 (1966).

Sugiura, A., Tobita, K., Kilbourne, E. D.: Isolation and preliminary characterization of temperature-sensitive mutants of influenza virus. J. Virol. *10,* 639–647 (1972).

Sugiura, A., Ueda, M.: Marker rescue with ultraviolet-inactivated influenza virus. J. Virol. *7,* 499–503 (1971).

Sugiura, A., Ueda, M., Tobita, K., Enomoto, C.: Further isolation and characterization of temperature-sensitive mutants of influenza virus. Virology *65,* 363–373 (1975).

Tobita, K., Kilbourne, E. D.: Structural polypeptides of antigenically distinct strains of influenza B virus. Arch. Virol. *47,* 367–374 (1975).

Tumova, B., Pereira, H. G.: Genetic interaction between influenza A viruses of human and animal origin. Virology *27,* 253–261 (1956).

Ward, C. W., Dopheide, T. A.: Evolution of the Hong Kong influenza A sub-type. Biochem. J. *195,* 337–340 (1981).

Webster, R. G., Campbell, C. H., Granoff, A.: The *in vivo* production of new influenza A viruses. I. Genetic recombination between avian and mammalian influenza viruses. Virology *44,* 317–328 (1971).

Webster, R. G., Campbell, C. H., Granoff, A.: The *in vivo* production of "new" influenza viruses. III. Isolation of recombinant influenza viruses under simulated conditions of natural transmission. Virology *51,* 149–162 (1973).

Webster, R. G., Laver, W. G.: Antigenic variation of influenza viruses. In: The Influenza Viruses and Influenza (Kilbourne, E. D., ed.), 270–314. New York: Academic Press 1975.

Webster, R. G., Laver, W. G., Air, G. M., Schild, G. C.: Molecular mechanisms of variation in influenza viruses. Nature *296,* 115–121 (1982).

Webster, R. G., Laver, W. G., Granoff, A.: Genetic reassortment with orthomyxoviruses. In: Virus Research (Fox, C. F., Robinsion, W. S., eds.), 513–524. New York: Academic Press 1973.

Wiley, D. C., Wilson, I. A., Skehel, J. J.: Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature *289,* 373–378 (1981).

Wilson, I. A., Skehel, J. J., Wiley, D. C.: Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature *289,* 366–373 (1981).

Young, J. F., Desselberger, U., Palese, P.: Evolution of human influenza A viruses in nature. Sequential mutations in the genomes of new H1N1 isolates. Cell *18,* 73–83 (1979).

Young, J. F., Palese, P.: Evolution of human influenza A viruses in nature: Recombination contributes to genetic variation of H1N1 strains. Proc. Natl. Acad. Sci. U.S.A. *76,* 6547–6551 (1979).

Young, J. F., Taussig, R., Aaronson, R. P., Palese, P.: Advantages and limitations of the oligonucleotide mapping technique for the analysis of viral RNAs. In: Replication of Negative Strand Viruses (Bishop, D. H. L., Compans, R. W., eds.), 209–215. New York: Elsevier/North-Holland 1981.

2

The Influenza Virus RNA Segments and Their Encoded Proteins

R. A. Lamb

I. Introduction

In this chapter the structure of the genes of influenza virus and their encoded proteins will be described. Influenza viruses contain a segmented single-stranded RNA genome which has been called "negative stranded" because the viral messenger RNAs are transcribed from the viral RNA segments. A great deal of new knowledge has been obtained about influenza A, B, and C viruses since the last major reviews of influenza virus in 1975 (Kilbourne, 1975). The complete sequence of the 8 RNA segments of influenza A virus has been obtained and this has led to a greatly increased understanding of the encoded polypeptides and how variation among influenza viruses occurs. A large body of evidence concerning changes in the surface glycoproteins, the hemagglutinin and neuraminidase, has been obtained from studies on the primary sequences, interactions with antibodies and the correlation of these data with the three-dimensional X-ray structure of these proteins.

In the sections below the basic structure of the virus and the methods of determining the gene structure will be discussed, followed by a more detailed review of the structure of each RNA segment and its encoded protein.

II. The Influenza Virus Particle: Basic Structure

Early information on the structure of influenza virus was from electron micrographs. When repeatedly cultured in the chorio-allantoic membrane of chick embryo, influenza virions have a fairly regular appearance in the electron microscope with particles 80–120 nm in diameter when negatively stained. In contrast, virus strains isolated from man or animals and propagated by limited passages in culture exhibit greater heterogeneity and pleomorphism (Horne *et al.*, 1960; Hoyle *et al.*, 1961). The most striking feature of the virion is a lipid envelope bearing surface projections or spikes radiating outwards. These spikes are of two types: rod-shaped spikes of hemagglutinin (HA) and, in much lesser abundance, mushroom-

shaped spikes of neuraminidase (NA) (Laver and Valentine, 1969). Both HA and NA are antigenic glycoproteins, but HA is believed to be the major antigenic determinant of the virus. However, recurrent influenza epidemics are associated with changes in the antigenic structure of both HA and NA.

The HA is involved in attachment of the virus to cellular receptors and also in initiating the infection process. The role of the NA is less clear but it hydrolyzes sialic acid from hemagglutinin receptors on the cell surface and may allow the release of progeny virus from the cell membrane. Influenza virus particles form by a process of budding at the plasma membrane (Murphy and Bang, 1952; Morgan *et al.*, 1961; Compans and Dimmock, 1969; Bachi *et al.*, 1969; Klenk *et al.*, 1970), with the lipid membrane of the virus, into which the surface spikes are inserted, being derived from the infected cell plasma membrane (Klenk and Choppin, 1969, 1970 a, b). Within the lipid envelope is an electron-dense layer composed of the viral membrane protein (M_1). This protein is believed to stabilize the structure of the virus particle. M_1 may also be an organizer for assembly of the virus at the plasma membrane before budding and maturation of the virus particle.

Inside the virus, observable by thin-sectioning of virus or by disrupting particles, are the ribonucleoprotein (RNP) structures which contain 8 different seg-

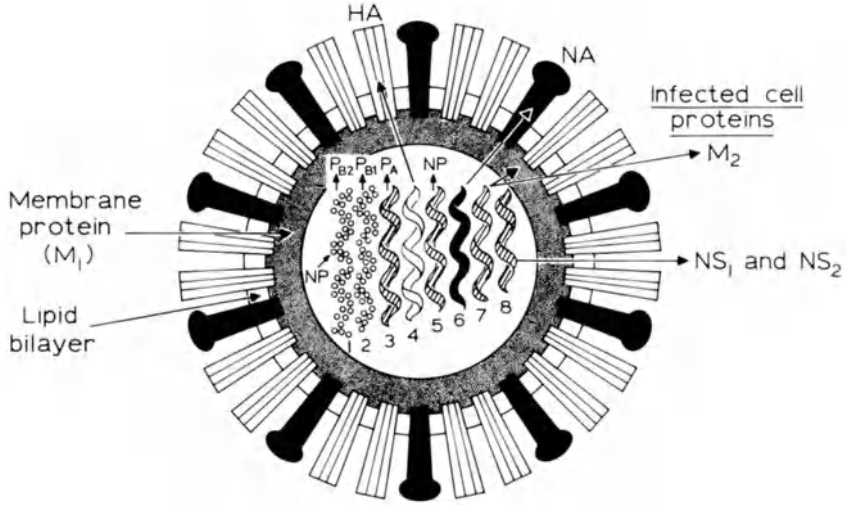


Fig. 1. A schematic diagram of the structure of the influenza virion. Two types of glycoprotein spikes, which possess hemagglutinating (HA) and neuraminidase *(NA)* activities, are inserted through the lipid bilayer of the viral membrane. The distributions of the spikes are not meant to reflect the much greater abundance of these proteins in the virion. On the inner surface of the lipid bilayer is the membrane protein *(M_1)* which is the major structural protein of the viral membrane. Within the envelope are the eight segments of single-stranded genome RNA contained in the form of helical ribonucleoproteins. Segments 1 and 2 are drawn to show the nucleocapsid *(NP)* protein associated with the RNA to form the ribonucleoprotein complex, which together with small amounts of proteins, PB1, PB2, and PA, has RNA-dependent RNA polymerase activity. The coding assignments of the 8 RNA segments are also illustrated. RNA segments 7 and 8 each code for more than one protein; however M_2, NS_1 and NS_2 are only found in infected cells and are not structural components of the virus. The diagram is not drawn to scale

ments of single-stranded RNA. The RNPs have the appearance of flexible rods (Pons *et al.*, 1969); depending on the salt concentration during sample preparation or method of staining, the flexibility appears to vary (Pons *et al.*, 1969; Compans *et al.*, 1972; Heggeness *et al.*, 1982). The RNP strands sometimes exhibit loops on one end and a periodicity of alternating major and minor grooves suggesting that the structure is formed by a strand that is folded back on itself and then coiled into a double helix. Since circular forms have also been observed the termini of the RNA may not necessarily be at one end of the rod (Heggeness *et al.*, 1982). The RNPs can be separated into different size classes (Duesberg, 1969; Pons, 1971; Rees and Dimmock, 1981) and it is thought that each RNP contains one segment of RNA. Larger helical structures, interpreted to be formed by the orderly aggregation of the smaller segments, have occasionally been seen in partially disrupted virions (Apostolov and Flewett, 1965; Murti *et al.*, 1980; Hoyle *et al.*, 1961; Almeida and Brand, 1975). These structures might represent ordered complexes of the 8 separate RNPs.

The RNPs consist of 4 protein components and RNA. The nucleoprotein (NP) is the predominant protein subunit of the nucleocapsid and it can be calculated that one NP molecule interacts with approximately 20 nucleotides. Also associated with the RNPs are three proteins found in lesser quantities, designated PB1, PB2 and PA. The RNP complex has RNA-dependent transcriptase activity (Bishop *et al.*, 1972; Inglis *et al.*, 1976; Compans and Caliguiri, 1973; Rochovansky, 1976; Plotch *et al.*, 1981). A schematic diagram of the structure of the virion is shown in Fig. 1.

III. Structure of the Genome

The genetic information of influenza virus is 8 segments of single-stranded RNA. These code for the 7 proteins found in the virion (PB1, PB2, PA, HA, NA, NP, and M1), which have both structural and enzymatic functions, and also for three proteins synthesized only in infected cells (the non-structural proteins NS_1, NS_2, and M_2), whose functions are not yet understood. The RNA of the virus is not infectious and the mRNAs from which the proteins are translated are transcribed from the virion RNA by the virion-associated RNA polymerase (transcriptase); some of these mRNAs are subsequently processed in the cell, most likely in the cell nucleus. Hence, the term "negative-stranded" has been adopted to describe influenza virus RNA as, by convention, mRNA is plus-stranded (Baltimore, 1971). The complete nucleotide sequence of influenza virus has now been established and considerable information is known about the structure of the proteins, including the X-ray crystallographic structure to 3 Å resolution of HA and NA. The evidence on which these conclusions are based will now be considered.

A. Early Evidence for a Segmented Genome

The genome of influenza virus was found to consist of single-stranded RNA (Ada and Perry, 1954; Duesberg and Robinson, 1967). Burnet and Lind (1951, 1954) found

a high rate of recombination of selectable markers between two strains, and with the introduction of plaque assays this could be confirmed for single virus isolates (Simpson and Hirst, 1961). The phenomenon of multiplicity reactivation after subjecting virus to inactivating dosages of U.V. radiation led Barry (1961) to propose that the genome consisted of about six independent genetic units. The high frequency of recombination (reassortment) led Hirst (1962) to propose that the genome was segmented. Further indirect evidence was obtained from studies using chemical inactivation of virus (Scholtissek and Rott, 1964). Biochemical characterization of the virion RNA by sucrose density sedimentation supported the concept that the genome was segmented (Davies and Barry, 1966; Duesberg and Robinson, 1967) and analysis by polyacrylamide gel electrophoresis indicated about 6 species of RNA (Duesberg, 1968; Pons and Hirst, 1968; Skehel, 1971; Bishop *et al.*, 1971; Choppin and Pons, 1970; Horst *et al.*, 1972). Evidence suggesting that the RNA segments were independently replicated came from finding that all segments had a 5′ terminus of pppA and a 3′ terminus of U-OH (Lewandowski *et al.*, 1971; Young and Content, 1971). Further genetic evidence for a segmented genome was provided by the isolation of temperature sensitive mutants (Simpson and Hirst, 1968; Mackenzie, 1970; Mills and Chanock, 1971; Ueda, 1972, Sugiura *et al.*, 1972; Ghendon *et al.*, 1973; Scholtissek *et al.*, 1974; Almond *et al.*, 1977) among which a high frequency of recombination was observed and which could be placed in non-overlapping complementation groups. It should be emphasized that recombination as used to describe an exchange of genetic information in influenza virus normally occurs by reassortment of separate genome RNA segments in a mixed infection.

B. The 8 Segments of Influenza Virus RNA

A major step forward in understanding the structure of the influenza virus genome was the electrophoretic separation of the virion RNA segments on polyacrylamide gels containing 6 M urea, which was adapted from a procedure developed for reovirus RNA (Floyd *et al.*, 1974). It was found that influenza A virus RNA could be separated into 8 segments (Bean and Simpson, 1976; Pons, 1976; Palese and Schulman, 1976 a, b; Ritchey *et al.*, 1976 a; McGeoch *et al.*, 1976, see Fig. 2) which are now known to range from 890—2341 nucleotides in length. The important demonstration that the eight species of RNA are distinct was done by two-dimensional oligonucleotide fingerprinting (McGeoch *et al.*, 1976). The segments have been numbered 1 to 8 in decreasing order of their electrophoretic mobility in polyacrylamide gels. It seemed likely that each RNA segment would code for one protein, since the approximate coding capacity of each RNA segment, estimated from its chain length, corresponded with the sizes of the known influenza virus proteins (Palese and Schulman, 1976 a; McGeoch *et al.*, 1976; Inglis *et al.*, 1976; Lamb and Choppin, 1976; Ritchey *et al.*, 1976 b). The RNA segments from different strains of influenza virus had differing migration patterns in acrylamide/urea gels (Palese and Schulman, 1976 a, see Fig. 2). Such differences have been exploited in assigning gene functions in recombination experiments (see below). However, when the RNA segments were completely denatured with glyoxal (McMaster and

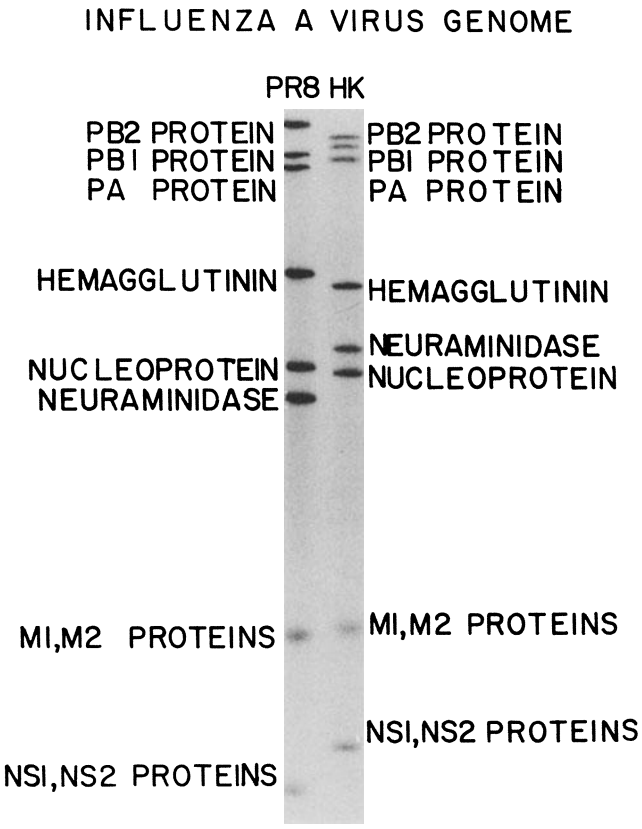


Fig. 2. Map of the influenza A virus genome. The RNAs of influenza A/PR/8/34 and A/HK/8/68 viruses (PR8 and HK viruses) were separated on a 2.6% polyacrylamide gel containing 0.15% N,N′-methylenebisacrylamide. Migration of the RNAs is from top to bottom. A gene product(s) has (have) been assigned to each of the eight influenza A virus RNA segments (see text). It should be noted that migration of the RNAs under the conditions used is dependent on the size as well as on the secondary structure of the molecules. Modified from Ritchey *et al.* (1976 b)

Carmichael, 1977), the mobilities of the RNA segments were identical between strains, with the exception of the RNA segments coding for HA and NA, which varied among influenza A viruses belonging to different HA and NA subtypes (Desselberger and Palese, 1978). Thus, the migration differences observed between the other RNA segments in acrylamide/urea gels are probably fortuitous and dependent on secondary structure.

The RNAs of influenza B and C viruses have also been examined by the methods described above. Influenza B virus is now known to contain 8 RNA segments and influenza C virus has been found to contain 7 RNA segments with the B virus RNA segments ranging from approximately 1096–2500 nucleotides in length and the C virus segments ranging from 1000–3000 nucleotides in length (Ritchey *et al.*, 1976 a; Cox and Kendal, 1976; Desselberger and Palese, 1978; Compans *et al.*,

1977; Racaniello and Palese, 1979; Briedis *et al.*, 1981). The inability to find 8 RNA segments in influenza C virus may reflect the presence of only one glycoprotein in the virions, since they lack neuraminidase activity (Kendal, 1975; Nerome *et al.*, 1978; Meir-Ewert *et al.*, 1978). It is known that influenza C viruses use a different receptor on erythrocytes (Hirst, 1950; Kendal, 1975; Meier-Ewert *et al.*, 1978) and contain sialic acid on the virus hemagglutinin (Nakamura *et al.*, 1979).

C. Methods for Assigning Gene Functions

The assignment of specific RNA segments to virus polypeptides has been made in three ways. The first method depended on the apparent differences between strains in the electrophoretic mobilities of RNA segments in polyacrylamide/urea gels and the ability to distinguish proteins between strains, either by immunological methods for the surface glycoproteins or by differences in the electrophoretic mobilities of the polypeptides in protein gels. Recombinants between two strains of virus were prepared and comparisons were made of the mobilities of the RNA segments or polypeptides between the parental types and each recombinant. From these analyse, gene assignments could be made (Palese and Schulman, 1976 b; Schulman and Palese, 1976; Ritchey *et al.*, 1976 b, 1977; Palese *et al.*, 1977 a; Lamb and Choppin, 1979, 1981). A similar approach was also used to assign a function to each of the complementation groups of temperature sensitive (ts) mutants (Almond *et al.*, 1977, 1979; Palese *et al.*, 1977 b).

The second method depended on base sequence homologies between corresponding RNA segments of different strains. Base sequence homologies were calculated between individual RNA segments of the A/FPV strain, which can plaque on chick embryo cells, and the complementary RNA segments of various other influenza virus strains which were unable to plaque on these cells. Temperature sensitive mutants of A/FPV with a known phenotypic defect were used to form recombinants with a non-plaquing strain. Rescued recombinant plaques at the non-permissive temperature were used to infect cells and complementary RNA was extracted and used in molecular hybridization studies with ^{32}P-labeled vRNA segments. By comparing the ribonuclease-resistant fraction of the homologous hybrids with that of the heterologous hybrid, the parental origin of any segment could be defined (Scholtissek *et al.*, 1976; Rohde *et al.*, 1977, 1978).

The third method is the most direct and is based on the ability to translate influenza virus mRNAs *in vitro* using cell extracts but the inability to translate a mRNA/vRNA hybrid. Cytoplasmic poly(A)-containing mRNA molecules were extracted from infected cells and hybridized to isolated vRNA segments. After hybridization, the total RNA was translated in wheat germ extracts. Each vRNA segment specifically prevented the translation of the polypeptide for which it codes (Inglis *et al.*, 1977; Lamb and Choppin, 1979, 1981; Inglis *et al.*, 1979) (see Fig. 3).

Detailed interpretations of some of these experiments are complicated by the fact that corresponding RNA segments of different strains do not migrate in the same order on acrylamide/urea gels, *e.g.*, segment 5 codes for NP in some strains and NA in others (Palese and Schulman, 1976 b; Almond *et al.*, 1977). However, as dis-

cussed above, after glyoxal treatment of RNA the segments from all strains migrate in the same order (Desselberger and Palese, 1978). It was also discovered that corresponding P proteins do not migrate in the same order on polyacrylamide gels, *e.g.*, the originally designated P2 protein of one strain can be the functional equivalent of the originally designated P3 protein of another strain (Almond *et al.*, 1979). For this reason the nomenclature of the P proteins, as based on the order in which they migrate on polyacrylamide gels has been changed and now reflects the RNA segment of origin (see below).

It is now accepted that by numbering the RNA segments according to their migration order on glyoxal gels (which with some strains is the same as acrylamide/urea gels) the gene assignments are as follows:
RNA segment 1 codes for PB2, 2 for PB1, 3 for PA, 4 for HA, 5 for NP, 6 for NA, 7 for M_1 and M_2 and 8 for NS_1 and NS_2.
A summary of these gene assignments is shown in Table 1.

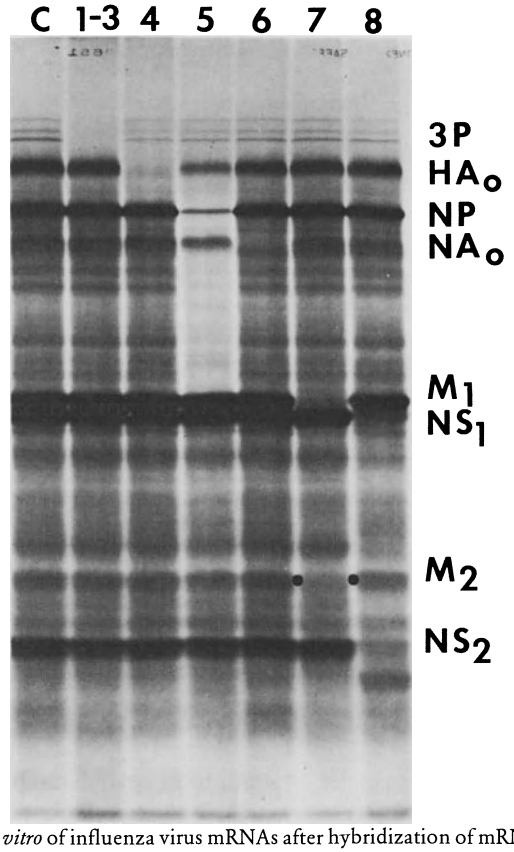


Fig. 3. Translation *in vitro* of influenza virus mRNAs after hybridization of mRNA from infected cells to individual genome RNA segments. *C* control *in vitro* translation of influenza virus mRNAs. Lanes *1–3, 4, 5, 6, 7, 8* show the polypeptides synthesized after hybridization of the respective virion RNA segments to mRNA. Note that RNA segment 7 inhibits the translation of M_1 and M_2 and RNA segment 8 inhibits the translation of NS_1 and NS_2 (modified from Lamb and Choppin, 1979)

Table 1. *Influenza virus genome RNA segments and coding assignments*

Segment	Length[a] (nucleotides)	mRNA length[b] (nucleotides)	Encoded[c] polypeptide	Nascent[d] polypeptide length (aa)	Approx. No.[e] molecules per virion	Remarks[f]
1	2,341	2320	PB2	759	30–60	Host cell RNA cap binding: component of RNA transcriptase
2	2,341	2320	PB1	757	30–60	Initiation of transcription: possibly endonuclease activity: component of RNA transcriptase
3	2,233	2211	PA	716	30–60	Elongation of mRNA chains?: component of RNA transcriptase
4	1,778	1757	HA	566	500	Surface glycoprotein, trimer; major antigenic determinant
5	1,565	1540	NP	498	1000	Associated with RNA segments to form ribonucleoprotein: structural component of RNA transcriptase
6	1,413	1392	NA	454	100	Surface glycoprotein: neuraminidase activity. Tetramer
7	1,027	1005	M_1	252	3000	Major protein component of virus: underlies lipid bilayer
		316	M_2	96		Spliced mRNA, non-structural protein: function unknown
		276	?	?9		Spliced mRNA, peptide predicted by nucleotide sequence only
8	890	868	NS_1	230		Non-structural protein: function unknown
		395	NS_2	121		Spliced mRNA, non-structural protein: function unknown
	13,588					

[a] For A/PR/8/34 strain: see Table 2 for references.
[b] Deduced from RNA sequence, excluding polyA tract.
[c] Determined by biochemical and genetic approaches (see text).
[d] Determined by nucleotide sequence analysis and protein sequencing (see text).
[e] Adapted from Compans and Choppin (1975).
[f] [illegible]

D. Sequences at the 5′ and 3′ Ends of Each RNA Segment Are Common

The terminal nucleotide sequences of the individual RNA segments of influenza virus are of considerable interest because of their common involvement in transcription, translation and replication. Limited direct RNA sequencing of the 5′- and 3′-ends of the 8 segments of RNA from influenza A viruses showed the presence of a common sequence of 13 nucleotides at the 5′ terminus of each segment and a common sequence of 12 nucleotides at the 3′ terminus of each segment (Skehel and Hay, 1978; Robertson, 1979; Desselberger *et al.*, 1980) (with some strains there is either a C or U residue at nucleotide 4 at the 3′ end). In addition, the conserved 3′- and 5′-terminal sequences show partial and inverted complementarity which may be important for replication of the RNA segments (Robertson, 1979; Desselberger *et al.*, 1980). The RNA-dependent polymerase (transcriptase) may recognize the 3′-end sequences of the virion (−) strand and because of the inverted complementary sequences at the 5′ and of the same strand, the polymerase could then bind to similar sequences at the 3′ end of the complementary (+) strand RNA during replication. A third common region in all 8 RNA segments is a uridine tract 15–21 residues from the 5′ end of the genome segments. This is now known to be the signal for initiating polyadenylation during mRNA synthesis (Robertson *et al.*, 1981) (see below). The RNAs of influenza B and C viruses also contain common 3′ and common 5′ ends and these are similar to the termini of the influenza A viruses (Skehel and Hay, 1978; Desselberger *et al.*, 1980).

E. Synthesis of Double-Stranded DNA from Influenza Virus RNA, Cloning and Nucleotide Sequencing

Although the total nucleotide sequence of influenza virus RNA segments could have been determined by successive ribonuclease digestions and homochromatography (Sanger *et al.*, 1965), only limited progress was made (Smith *et al.*, 1978) due to the laboriousness of the method. The development of molecular cloning of double-stranded DNA derived from single-stranded RNA, coupled with rapid DNA sequencing methods (Maxam and Gilbert, 1977; Sanger *et al.*, 1977) made sequencing the influenza virus genome much easier. The availability of purified virion RNA (vRNA) made it the preferred template from which to make cDNA copies rather than mRNAs isolated from infected cells (especially as the mRNAs are incomplete copies of the vRNA segment). To synthesize cDNA copies of vRNA with reverse transcriptase a primer is necessary. Many workers added poly(A) tails to the 3′ ends of the vRNA segments using polyadenyltransferase isolated from *E. coli* (Sippel, 1973) and then primed the reverse transcriptase with either oligo (dT) or oligo (dT-dA) (a U is the 3′ nucleotide of the vRNA strand) (Emtage *et al.*, 1979; Sleigh *et al.*, 1979; Caton and Robertson, 1979; Porter *et al.*, 1979; Both and Air, 1979; Verhoeyen *et al.*, 1980; Gething *et al.*, 1980; Winter and Fields, 1980; Briedis and Lamb, 1982). Some workers made the cDNA copy double-stranded (ds) by self-priming and using either the Klenow fragment of *E. coli* DNA polymerase 1 or

reverse transcriptase. The disadvantage of this method is that in order to clone the ds DNA in a cloning vector, a digestion with nuclease S1 has to be used and some sequences are lost from the cDNA molecule (Porter *et al.*, 1979; Verhoeyen *et al.*, 1980; Gething *et al.*, 1980). An alternative to priming with oligo dT was to use a synthetic oligonucleotide dodecamer d[A-G-C-A-A-A-A-G-C-A-G-G], which is complementary to the 3′ end of every virion RNA segment (Lai *et al.*, 1980; Baez *et al.*, 1980; Winter *et al.*, 1981 b). Another approach to making ds DNA was based on a procedure developed for cloning bacterophage Qβ (Taniguchi *et al.*, 1978). A (+) cDNA copy was transcribed from the vRNA and a (−) cDNA copy was obtained from mRNA molecules which were isolated from infected cells. The (+) and (−) cDNA copies were hybridized and then inserted into a cloning vector (Caton and Robertson, 1979, 1980; Lai *et al.*, 1980) (see Fig. 4). The advantage of this method is that complete cDNA copies of both the vRNA and mRNA are more readily obtained (Lamb and Lai, 1980, 1981; Dhar *et al.*, 1980; Markoff and Lai, 1982; McCauley *et al.*, 1982; Briedis and Lamb, 1982; Briedis *et al.*, 1982; Shaw *et al.*, 1982; Sveda *et al.*, 1982). Another strategy used the dodecanucleotide primer complementary to the 3′ end of vRNA segments to synthesize a (+) cDNA copy and then another synthetic oligonucleotide complementary to the 3′ end of the (+) cDNA copy for synthesis of the second DNA strand. The use of two synthetic primers has enabled complete dsDNA copies of each vRNA segment to be obtained (Baez *et al.*, 1981; Winter *et al.*, 1981 b; Lamb and Lai, 1982; Lin and Lai, 1982; Bishop *et al.*, 1982 a, b).

A variety of cloning vectors have been used for amplifying the dsDNA in bacteria, including the plasmid pBR322 and its high-copy derivative pAT153, bacteriophage lambda and the replicative form of derivatives of bacteriophage M13 (M13mp2 and M13mp7) (Table 2).

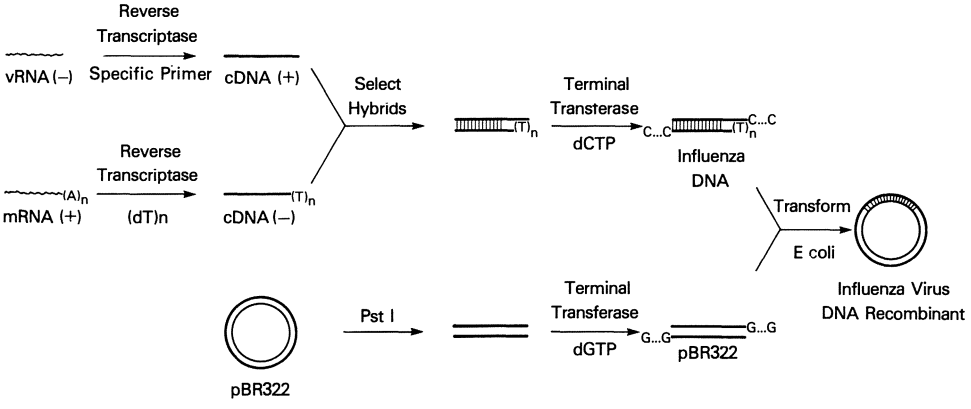


Fig. 4. Scheme for construction and cloning of influenza virus DNA. cDNA copies of both virion RNA and mRNA strands are made and then hybridized to form double-stranded DNA. The advantage of this method is that complete copies of both the vRNA and mRNA are cloned. (From Lai *et al.*, 1980 and kindly provided by Dr. Ching-Juh Lai)

Table 2. *Compilation of complete nucleotide sequences of influenza virus RNA segments*

RNA segment[a]	Encoded protein	Strain		Authors
1	PB2	A/PR/8/34		Fields and Winter (1982)
		A/WSN/33		Kaptein and Nayak (1982)
2	PB1	A/PR/8/34		Winter and Fields (1982)
		A/WSN/33		Sivasubramanian and Nayak (1982)
		A/NT/60/68		Bishop *et al.* (1982 a)
3	PA	A/PR/8/34		Fields and Winter (1982)
		A/NT/60/68		Bishop *et al.* (1982 b)
4	HA	A/PR/8/34	H1	Winter *et al.* (1981 a)
		A/WSN/33	H1	Hiti *et al.* (1981)
		A/Japan/305/57	H2	Gething *et al.* (1980)
		A/Duck/Ukraine/63	H3	Fang *et al.* (1981)
		A/Aichi/2/68	H3	Verhoeyen *et al.* (1980)
		A/NT/60/68	H3	Both and Sleigh (1980)
		A/Memphis/102/72	H3	Sleigh *et al.* (1980)
		A/Victoria/3/75	H3	Min Jou *et al.* (1980)
		A/Bangkok/1/79	H3	Both and Sleigh (1981)
		A/FPV/Rostock/34	H7	Porter *et al.* (1979)
		B/Lee/40		Krystal *et al.* (1982)
5	NP	A/PR/8/34		Winter and Fields (1981)
		A/PR/8/34		Van Romuy *et al.* (1981)
		A/NP/60/68		Huddleston and Brownlee (1982)
6	NA	A/PR/8/34	N1	Fields *et al.* (1981)
		A/WSN/33	N1	Hiti and Nayak (1982)
		A/NT/60/68	N2	Bentley and Brownlee (1982)
		A/Udorn/72	N2	Markoff and Lai (1982)
	NA and NB	B/Lee/40		Shaw *et al.* (1982) Shaw and Lamb (unpublished results)
7	M_1 and M_2	A/PR/8/34		Winter and Fields (1980)
		A/PR/8/34		Allen *et al.* (1980)
		A/Udorn/72		Lamb and Lai (1981)
		A/FPV/Rostock/34		McCauley *et al.* (1982)
		B/Lee/40		Briedis *et al.* (1982)
8	NS_1 and NS_2	A/PR/8/34		Winter *et al.* (1981 b)
		A/PR/8/34		Baez *et al.* (1980)
		A/Udorn/72		Lamb and Lai (1980)
		A/Duck/Alberta/60/76		Baez *et al.* (1981)
		A/FPV/Rostock/34		Porter *et al.* (1980)
		B/Lee/40		Briedis and Lamb (1982)

[a] In migration order on polyacrylamide gels after glyoxal treatment (Desselberger and Palese, 1978).

Nucleotide sequencing has been done by the partial chemical degradation method of Maxam and Gilbert (1977) or for M13 cloning by the dideoxy-substitution chain-terminating sequencing method (Sanger *et al.*, 1977). Originally Winter and Fields (1980) expected to sequence the entire influenza virus genome using the latter technique following the "shot-gun" cloning into M13 of restriction endo-

nuclease-digested pieces of dsDNA (Gronenborn and Messing, 1978). Such a strategy was developed for sequencing mitochondrial DNA (Sanger *et al.*, 1980). However, it became apparent that not all influenza virus sequences were cloned equally easily and gaps in the sequence had to be covered by dideoxy sequencing using as primers restriction fragments of existing M13 clones. Some of the gaps in the sequences can be attributed to *EcoK* sites in segments 5 and 7 at these points (Winter and Fields, 1982). Subsequently, it also became easier to clone into pBR322 the complete dsDNA which was synthesized from vRNA, using two oligonucleotide primers, and then to subclone into M13 before sequencing (Winter *et al.*, 1981; Winter and Fields, 1982). From such strategies the complete sequence of strain A/PR/8/34 was obtained (see Table 2 for references).

Considerable information has been obtained, particularly for different HA and NA subtypes, by modifying the dideoxy chain-termination method so that reverse transcriptase primed with oligo (dT_8-dA) copies a polyadenylated vRNA segment. The nucleotide sequencing gels could then be done directly from this reaction and hence it is a very rapid method. One disadvantage of this technique is that as priming only takes place at the 3′ end of the vRNA strand, the length of nucleotide sequence obtained is related to the resolution of the polyacrylamide gel used to separate oligonucleotides of differing lengths (Air, 1979, 1981; Both and Air, 1979; Air and Hackett, 1980; Block and Air, 1980, 1982; Block *et al.*, 1982). A modification of this scheme was used to obtain complete sequences of variants of HA. Restriction endonuclease fragments from a cloned DNA copy of HA were used to prime on vRNA of other viruses (Both and Sleigh, 1980; Sleigh *et al.*, 1980).

IV. The Influenza Virus RNA Segments

The complete nucleotide sequence of all 8 RNA segments of an influenza A virus has now been obtained (Table 2) and segments 4, 6, 7, and 8 of an influenza B virus have also been sequenced. The structure of each RNA segment and properties of its encoded polypeptide(s) are described in the sections below. No sequence information concerning influenza C virus RNAs is yet available.

A. RNA Segments 1, 2, and 3: Properties of the Transcriptase Associated Proteins PB1, PB2, and PA

In early studies, either one or two "P" proteins were found to be associated with the ribonucleoprotein complex of virions (Compans *et al.*, 1970; Skehel and Schild, 1971; Skehel, 1972; Bishop *et al.*, 1972). By using improved polyacrylamide gel electrophoresis conditions it became clear that the virion contained three discrete proteins (Lamb and Choppin, 1976; Inglis *et al.*, 1976; Plotch *et al.*, 1981) which were designated P1, P2, P3 in decreasing order of their electrophoretic mobility on polyacrylamide gels (apparent MW 95–82,000). From analysis of recombinants of strains A/PR/8/34 and A/HK/68 it was determined that RNA segment 1 coded for the P3 protein, RNA segment 2 for the P1 protein and RNA segment 3 for the P2

protein (Palese *et al.*, 1977 a; Ritchey *et al.*, 1977) but with A/FPV/Rostock/34 the coding assignment to RNA segments was different (Scholtissek *et al.*, 1976). Since the assignment of the nomenclature P1–P3 is based on gel mobility and not function, the P2 protein of one strain can be the functional equivalent of P3 of another strain (Almond and Barry, 1979).

A better identification of the P proteins has been achieved by using a 2-dimensional electrophoresis system with isoelectric focusing of the proteins in the first dimension and SDS electrophoresis in the presence of reducing agents in the second dimension (Horisberger, 1980). It was found that two of the P polypeptides were relatively basic and one of the P polypeptides was relatively acidic. This finding makes for a better nomenclature of the P proteins until all their functions are elucidated (Horisberger, 1980; Ulmanen *et al.*, 1981; Winter and Fields, 1982). The more rapidly migrating basic P protein coded by RNA segment 1 is called PB2, the slowly migrating basic P protein coded by RNA segment 2 is called PB1 and the acid P protein coded by RNA segment 3 is called PA (see Tables 1 and 3).

Table 3. *RNA segments 1, 2, 3 and their encoded polypeptides*[a]

Segment	Length nucleotides	Length of polypeptide amino acids	Mol. wt. of encoded polypeptide	Net change at pH 6.5	Gene	Assignment	New[d] nomenclature
					PR8[b]	FPV[c]	
1	2341	759	85,900	+28	P3	P2	PB2
2	2341	757	86,500	+28	P1	P1	PB1
3	2233	716	82,400	−13.5	P2	P3	PA

[a] Adapted from Winter and Fields (1982). Data from Fields and Winter (1981) and Winter and Fields (1982).
[b] Palese *et al.* (1977).
[c] Almond and Barry (1979).
[d] Based on data from Horisberger (1980), Ulmanen *et al.* (1981), Winter and Fields (1982).

The complete sequence of RNA segments, 1, 2, and 3 has been obtained (Fields and Winter, 1982; Winter and Fields, 1982; Bishop *et al.*, 1982 a, b; Kaptein and Nayak, 1982; Sivasubramanian and Nayak, 1982). RNA segments 1 and 2 are both 2341 nucleotides in length and code for proteins of 759 and 757 amino acids respectively. Both are basic proteins at pH 6.5 with a net charge of +28. As the sizes of the two polypeptides PB1 and PB2 are very similar, the ability to separate them on polyacrylamide gels must be due to factors such as differential binding of SDS. RNA segment 3 is 2233 nucleotides in length and codes for a protein of 716 amino acids that has a charge of −13.5 at pH 6.5. It has been suggested that, as the P proteins are involved in RNA synthesis and as they are so similar in size, these molecules evolved from a single ancestral polymerase gene. However, no simple homology could be detected from a comparison of their sequences (Winter and Fields, 1982).

It has been suggested from studies with *ts* mutants defective in RNA synthesis that PB1 and PB2 are involved in mRNA synthesis and PA in vRNA synthesis (Krug *et al.*, 1975; Palese *et al.*, 1977 b; Scholtissek and Bowles, 1975). Detailed analyses of

the *in vitro* transcription reaction indicate that PB2 is involved in binding to the cap-structure of the host-cell primer RNA (Ulmanen *et al.*, 1981; Blass *et al.*, 1982; Penn *et al.*, 1982; Ulmanen *et al.*, 1983) and PBI in initiating transcription (Ulmanen *et al.*, 1983).

B. RNA Segment 4: Structure and Function of the Hemagglutinin

The influenza virus hemagglutinin (HA) is an integral membrane glycoprotein responsible for the attachment of the virus to cells. It was originally named because of the ability of the virus to agglutinate erythrocytes (Hirst, 1941; McClelland and Hare, 1941) by attachment to specific sialic acid containing glycoprotein receptors (Hirst, 1942). The HA constitutes a large percentage of the visible spikes on the virion. It is the major antigen of the virus against which neutralizing antibodies are directed (Laver and Kilbourne, 1966) and recurrent influenza epidemics are associated with changes in its antigenic structure. In addition to mediating the attachment of infectious virus to the plasma membrane of the susceptible host cell, the HA is also responsible for initiating infection (Klenk *et al.*, 1975; Lazarowitz and Choppin, 1975). Depending on the virus strain, host cell type and growth conditions, the HA (MW ~ 77,000) can be cleaved proteolytically into two polypeptide chains, HA1 (MW ~ 50,000) and HA2 (MW ~ 27,000) (Lazarowitz *et al.*, 1971, 1973 a, b; Klenk *et al.*, 1972; Skehel, 1972), which are held together by disulfide bonds (Laver, 1971). This cleavage of HA does not affect its antigenic or receptor binding properties (Lazarowitz *et al.*, 1971, 1973 a, b; Lohmeyer and Klenk, 1979; McCauley *et al.*, 1980) but greatly increases the infectivity of the virus (Lazarowitz and Choppin, 1975; Klenk *et al.*, 1975) and is important for pathogenicity (Bosch *et al.*, 1979; Rott, 1979) and the spread of infection in the organism (Rott *et al.*, 1980).

Examination of HA by electron microscopy (Laver and Valentine, 1969; Griffith, 1975; Wrigley *et al.*, 1977) after its removal from the virion by detergent or selective proteolysis (Laver, 1964; Schulze, 1970; Compans *et al.*, 1970) showed that the molecule consisted of more than one subunit. It is now known to be a trimer of non-covalently linked monomers (Wiley *et al.*, 1977; Wilson *et al.*, 1981). Correlation of the polypeptide sizes of HA1 and HA2 isolated by protease release or detergent solubilization of virions, together with the observation that removal of the detergent from solubilized HA caused aggregation, whereas protease-released spikes were soluble, led to the suggestion that HA2 is inserted in the virus envelope (Laver and Valentine, 1969; Compans *et al.*, 1970; Brand and Skehel, 1972). This has been confirmed by protein sequencing studies showing that HA1 is the N-terminal portion of HA and HA2 the C-terminal portion (Skehel and Waterfield, 1975). Thus, the hydrophobic region of the molecule embedded in the membrane and removed by proteolysis was shown to reside in the C-terminal region of HA2.

1. The Structure of RNA Segment 4 Coding for the Hemagglutinin

The complete nucleotide sequence of RNA segment 4, coding for the HA, has been obtained for the H1, H2, H3, and H7 antigenic subtypes and for several variants of

Table 4. *Differences in the HA sequence of four subtypes*

	H1[a]	H2[b]	H3[c]	H7[d]
Nucleotide length	1778	1773	1765	1742
Coding region (nucleotides)	1698	1686	1698	1689
5′ untranslated region	32	43	29	21
3′ untranslated region	48	45	35	32
Coding region (aa)	566	562	566	563
Signal sequence (aa)	17	15	16	18
HA_1 (aa)	326	324	328	319
HA_2 (aa)	222	222	221	221
Amino acids lost between HA_1 and HA_2	1	1	1	5

[a] For A/PR/8/34 (Winter *et al.*, 1981). A/WSN/33 has one less amino acid in HA_1 and a nucleotide length of 1775 (Hiti *et al.*, 1981).

[b] For A/JAP/305/57 (Gething *et al.*, 1980). Substantial protein sequence data from Waterfield *et al.* (1979, 1980).

[c] For A/Aichi/2/68 (Verhoeyen *et al.*, 1980).
Several other HAs of the H3 subtype have been sequenced.
A/NT/60 (Both and Sleigh, 1980).
A/Victoria/75 (Min Jou *et al.*, 1980).
A/Bangkok/79 (Both and Sleigh, 1981).
A/Memphis/72 (Sleigh *et al.*, 1980).
Extensive protein sequencing of the A/Aichi/2/68 (X-31) and A/Memphis/72 was done (Dopheide and Ward, 1978; Ward and Dopheide, 1981; Ward, 1981).

[d] For A/FPV/Rostock/34 (Porter *et al.*, 1979).

the H1 and H3 strains (see Tables 2 and 4). Extensive direct protein sequence data has been obtained for the H2 and H3 strains (see reviews by Ward, 1981; Ward and Dopheide, 1979; Waterfield *et al.*, 1979). The overall basic structures of the different HA types is very similar. For details of the amino acid changes between the anti-

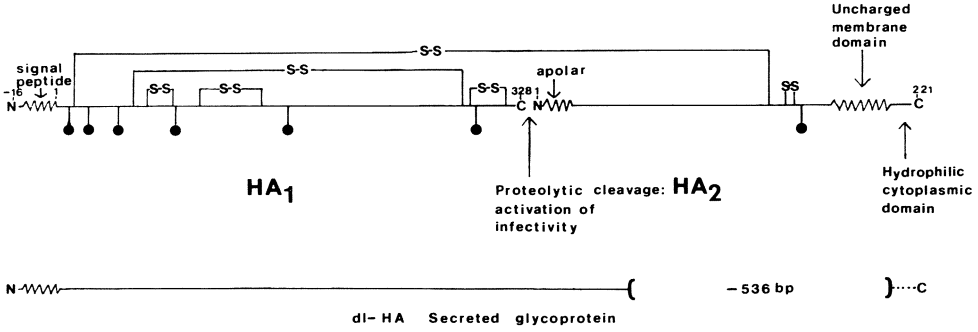


Fig. 5. Top: Schematic diagram of the A/Aichi/68 (H3) hemagglutinin polypeptide sequence. The N-terminal signal sequence, which is cleaved off during co-translational membrane insertion is shown. The position of the C-terminal uncharged membrane anchoring domain is shown, together with the site of proteolytic cleavage which yields polypeptide chains HA1 and HA2 and causes activation of infectivity. The disulfide bonds are shown and carbohydrate-attachment sites are illustrated by the line-and-circle. Adapted from Wilson *et al.* (1981). Bottom: Schematic diagram of a hemagglutinin that has lost a large portion of the C-terminus by genetic manipulation using a simian virus 40-HA expression vector (536 base pair deletion of DNA coding for HA). The synthesized HA molecule is a secreted glycoprotein. For details of this experiment see Sveda *et al.* (1982)

genically different HA subtypes and for changes within a subtype, the reader is referred to reviews by Ward (1981) and Webster *et al.* (1982), and this volume. The primary structure of the HA can be described using the A/Aichi/68 (H3) strain as an example (Verhoeyen *et al.*, 1980) as shown schematically in Fig. 5.

The RNA segment coding for HA is 1765 nucleotides in length: 20 nucleotides precede the signal for the AUG translation initiation codon, with the translated region extending over 1698 nucleotides (specifying 566 amino acids), before a 3′-terminal untranslated region comprised of 35 nucleotides. The hydrophobic signal peptide contains 16 amino acids; HA1 consists of 328 amino acids and HA2 has 221 amino acids. A single arginine residue is lost between HA1 and HA2 on proteolytic cleavage, suggesting that two enzymes are involved in the activation of the hemagglutinin: after the initial action of trypsin or a trypsin-like protease, an exopeptidase of the carboxypeptidase B type removes arginine from the cleavage site (Dopheide and Ward, 1978; Garten *et al.*, 1981). The N terminus of HA2 is hydrophobic and highly conserved between strains. Its possible role in interacting with the cell membrane is described below. Near the C terminus of HA2 is a stretch of 24 hydrophobic residues which span the lipid bilayer and anchor the HA in the lipid envelope. The precise location or function of the signal interrupting the translocation of the HA across the membrane ("Stop-Transfer Signal", Blobel, 1980) remain to be defined. However, this sequence must be within 178 amino acids of the C terminus of HA2, because, in a eukaryotic expression system containing cloned HA DNA from which this region was deleted the synthesis of a soluble HA molecule was observed (see Fig. 4) (Sveda *et al.*, 1982).

The carbohydrate composition of HA has been examined by a number of workers (Schwartz *et al.*, 1977; Keil *et al.*, 1979; Nakamura and Compans, 1978, 1979). In the A/Aichi/68 strain there are 7 possible glycosylation sites (asparagine-X-serine or asparagine-X-threonine). Analysis shows four complex carbohydrate chains on HA1 and one on HA2 consisting of N-acetylglucosamine, mannose, galactose and fucose and two simple carbohydrate chains on HA1, consisting of N-acetylglucosamine and mannose, all of which are linked via N-glycosidic linkages to asparagine residues (Ward and Dopheide, 1981; Wilson *et al.*, 1981). Terminal sialic acid residues are absent due to the action of the viral neuraminidase (Klenk *et al.*, 1970). There are 6 disulfide bonds in the HA molecule, 4 within the HA1 domain, one in HA2 and one linking HA1 and HA2 (Dopheide and Ward, 1980; Waterfield *et al.*, 1980).

2. *The Three-Dimensional Structure of the Hemagglutinin*

Cleavage of the A/Aichi/68 (H3) HA with bromelain releases an antigenically and structurally intact molecule which is water soluble and has lost only the C-terminal hydrophobic region of HA2 (Brand and Skehel, Skehel and Waterfield, 1975; Flanagan and Skehel, 1977; Wiley *et al.*, 1977; Waterfield *et al.*, 1979). This molecule can be crystalized and the three dimensional structure has been determined to 3 Å resolution by X-ray crystallography (Wilson *et al.*, 1981) (see Chapter 5, Fig. 4). The molecule is an elongated cylinder 135 Å long consisting of (a) a long fibrous stem extending 76 Å from the membrane containing two antiparallel α-helices which

terminate near the membrane in a compact five-stranded antiparallel β-sheet globular fold and (b) a distal globular region of antiparallel β-sheet structure. The latter globular region is composed entirely of residues from HA1 and is connected to the HA2 fibrous stem by only two antiparallel chains from HA1. The C terminus of HA1 is 21 Å from the N terminus of HA2, indicating a substantial rearrangement in this region when the HA is activated by cleavage. The N terminus of HA1 is very close to the membrane and it is interesting to speculate that the molecules may have folded while attached to the membrane not only by the hydrophobic C-terminal region of HA2 but also by the signal sequence at the N terminus of HA1, which was subsequently removed by proteolysis. The variable antigenic determinants are at the top of the globular region (see Chapter 5, Fig. 4). Four antigenic sites have been proposed (Wiley *et al.*, 1981; Webster and Laver, 1980; Gerhard *et al.*, 1980). These were derived from amino acid changes in antigenic variants selected with monoclonal antibodies (Laver, 1981) and in naturally arising variants of this virus (Wilson *et al.*, 1981; Laver *et al.*, 1980, 1981; Both and Sleigh, 1981; Yewdell *et al.*, 1979; Laver *et al.*, 1979). Further mapping of changes in amino acid sequence due to antigenic drift and correlation with the antigenic sites is continuing.

Although the X-ray structure has only been done on the H3 subtype, sequence comparisons with the H1 and H2 subtypes indicate that 22% of HA1 and 45% of HA2 residues are absolutely conserved, including the six disulfide bridges and many structurally important residues such as proline (21 residues) and glycine (9 residues). Therefore, it is expected that the three-dimensional structure of the HA of H1 and H2 virus subtypes is very similar to that of the H3 subtype.

Influenza B virus HA has also been sequenced from a cDNA copy of RNA segment 4 (Krystal *et al.*, 1982). The gene is 1882 nucleotides long and codes for 584 amino acids. The major structural features common to the A viruses are also found in the B virus HA, including a hydrophobic signal peptide, hydrophobic N and C termini of HA2 and the HA1/HA2 cleavage site. Over the length of the molecule, 24% of the amino acids in HA1 and 39% of the amino acids in HA2 are conserved between the A and B viruses, suggesting a close evolutionary relationship. The largest area of conserved sequences is the N terminus of HA2 which, as described above, is essential for the cleavage activation of infectivity.

3. Synthesis of the Hemagglutinin, Cotranslational and Post-Translational Modifications

The HA is translated as a single polypeptide chain from a monocistronic mRNA derived from virion RNA segment 4. The available data indicate that HA biosynthesis follows the same pathway as many cellular integral membrane proteins, with cotranslational insertion into the endoplasmic reticulum after attachment of the nascent polypeptide chain complexed with the ribosome (Blobel, 1980). The hydrophobic N-terminal signal peptide is removed by proteolysis (Air, 1979; McCauley *et al.*, 1979) and the elongating chain is glycosylated in the rough endoplasmic reticulum lumen (Compans, 1973 a; Hay, 1974; Klenk *et al.*, 1974; Elder *et al.*, 1979; McCauley *et al.*, 1980). Further glycosylation modifications occur during transport to the plasma membrane (Compans, 1973 b; Rott and Klenk, 1977). A fine dissection of the steps leading to the transport of the HA has not been reported, but

it is possible that clathrin-coated vesicles are involved as has been proposed for the vesicular stomatitis virus (VSV) glycoprotein (see Rothman *et al.*, 1982). However, in infected Madin-Darby bovine kidney cells, influenza virus buds from the apical surface of the cell, whereas VSV buds from its basolateral surface (Rodriguez-Boulan and Sabatini, 1978; Compans *et al.*, 1981; Roth *et al.*, 1979). The mechanism of sorting of these glycoproteins between cell surface remains to be determined.

4. Cleavage Activation of Infectivity and in Vitro Fusion

The most highly conserved sequence in the HA between strains is the apolar N terminus of HA2 which is associated with the activity by which the virus penetrates a host-cell membrane to initiate infection. Cleavage of the HA to HA1 and HA2 is required for infectious penetration and *in vitro* fusion (Lazarowitz *et al.*, 1975; Klenk *et al.*, 1975; Huang *et al.*, 1980) and results in a large conformational change in the HA (Skehel *et al.*, 1982). A sequence homologous to that at the N terminus of HA2 is present at the cleavage-activation site of the fusion protein (F) of Sendai virus (Scheid and Choppin, 1974, 1975; Gething *et al.*, 1978; Scheid *et al.*, 1978; Richardson *et al.*, 1980). The importance of this amino acid sequence to the fusion process is indicated by the finding that synthetic peptides mimicking the appropriate sequence of either the HA2 N terminus, or of the Sendai virus F_1 protein N terminus, inhibit virus penetration (Richardson *et al.*, 1980). However, there is a major difficulty in rationalizing how the N terminus of HA2 is directly involved in cell penetration, in as much as it is located 100 Å from the distal tip of the HA molecule and 35 Å from the membrane. Perhaps a second conformational change after receptor binding brings the HA2 terminus closer to the cell membrane. Influenza virus HA can cause cell fusion *in vitro* at acid pH (5.0—5.55) (Huang *et al.*, 1980, 1981; Maeda *et al.*, 1981; Lenard and Miller, 1981; White *et al.*, 1982; Matlin *et al.*, 1982) and it has been proposed that fusion activity at acid pH reflects entry of virus through the acidic environment of intracellular vesicles such as lysosomes.

C. RNA Segment 5: The Structure of the Nucleocapsid Protein (NP)

The nucleocapsid protein or nucleoprotein is encoded in RNA segment 5 and is the viral protein that interacts with the viral RNA segments to form the viral ribonucleoprotein (RNP) particles (Pons *et al.*, 1969; Compans *et al.*, 1972). The NP protein must be multifunctional as it presumably interacts not only with itself and RNA in the RNP particles but also with the three P (PB1, PB2, and PA) proteins to form the transcriptive complex. In addition, biochemical evidence has been obtained which indicates that the RNPs can interact with the membrane protein (M_1) (Rees and Dimmock, 1981) supporting the original concept that RNPs interact with the viral membrane protein at the plasma membrane (Choppin *et al.*, 1972). NP is the type specific antigen of influenza viruses used to distinguish the A, B, and C viruses, although minor differences in the antigenicity of NP among the A viruses have been observed (Schild *et al.*, 1979). Studies with monoclonal antibodies to NP

show that the molecule contains at least three non-overlapping antigenic areas and binding to one of these inhibits *in vitro* transcriptase activity (Van Wyke *et al.*, 1980). Whether this inhibition is by steric hinderance or interference with a function of NP other than its structural role in the complex is not clear. NP is phosphorylated at up to one serine residue per molecule (Kamuta and Watanabe, 1977; Privalsky and Penhoet, 1981; Petri and Dimmock, 1981) but it is not clear what percentage of NP molecules is phosphorylated or what role the phosphate group plays.

The complete sequence of RNA segment 5 has been obtained for two strains (A/PR/8/34; Winter and Fields, 1981; Van Rompuy *et al.*, 1981, with corrections by Min Jou, unpublished data 1981; A/NT/60/68; Huddleston and Brownlee, 1982). RNA segment 5 is 1565 nucleotides in length with, in the mRNA sense, a 5′ non-coding region of 45 nucleotides and a 3′ non-coding region of 26 nucleotides. The coding region of 1494 nucleotides codes for 498 amino acids which gives a predicted MW of 56,101 for NP which is very close to that deduced from its electrophoretic behavior (Pons *et al.*, 1969; Compans *et al.*, 1970; Lazarowitz *et al.*, 1971; Skehel, 1972). The protein is rich in arginine and has a net positive charge of +14 at pH 6.5. There are no clusters of basic residues which might have been predicted for the interaction of the acidic phosphate residues of RNA with the NP molecules, and therefore the RNA is probably associated with many regions of the NP molecule to neutralize the charges. Based on the total length of the influenza virus genome (13,588 nucleotides) and the number of molecules of NP associated with one virus particle (Compans and Choppin, 1975), it can be estimated that approximately 20 nucleotides interact with a single protein subunit. It is thought likely that the RNA is bound to the outside of the ribonucleoprotein structure, since it can be displaced by polyvinylsulfate (Pons *et al.*, 1969; Goldstein and Pons, 1970) and is susceptible to digestion with ribonuclease without disrupting the RNP structure (Duesberg, 1969; Kingsbury and Webster, 1969; Pons *et al.*, 1969; Murti *et al.*, 1980).

D. RNA Segment 6: The Structure and Properties of the Neuraminidase

RNA segment 6 codes for an integral membrane glycoprotein, the neuraminidase (NA). Influenza virions contain receptor-destroying neuraminidase activity which causes agglutinated red blood cells to disagglutinate; cells first treated with virus cannot be reagglutinated by fresh virus (Hirst, 1942). The neuraminidase hydrolytically cleaves the glycosidic bond joining the keto group of N-acetylneuraminic acid (sialic acid) to D-galactose or D-galactosamine (Gottschalk, 1957).

After NA is released from the virion membrane with detergent, electron microscopic examination reveals a mushroom-shaped structure with a stalk and head (Laver and Valentine, 1969; Wrigley *et al.*, 1973) consisting of a complex of four NA molecules with a total MW ~220,000. This model is confirmed by the X-ray three-dimensional structure (Coleman *et al.*, 1980; Varghese *et al.*, 1983). For a description of the 3 Å structure of NA see Chapter 5. The individual polypeptide chains (MW ~56,000) are held together by disulfide bonds (Lazdins *et al.*, 1972). Incomplete NA molecules can be isolated by protease digestion of virus particles;

these "head" molecules have lost their hydrophobic domains and their thin stalk while retaining enzymatic and antigenic properties (Mayron *et al.*, 1961; Noll *et al.*, 1962; Wilson and Rafelson, 1963; Lazdins *et al.*, 1972; Wrigley *et al.*, 1977; Blok *et al.*, 1982).

The function of NA in virus replication and maturation may be to promote release of budding virus particles from the host cell membrane (Palese *et al.*, 1974). The NA may also alter the host range of the virus by removing sialic acid and exposing HA to cleavage (Schulman and Palese, 1977) and through this mechanism NA has been implicated in aiding viral neurovirulence (Sugiura and Ueda, 1980). NA may also enable the virus to elute from inhibitory mucoproteins present in the respiratory mucosa allowing it to find a way to the target epithelial cells.

Antibodies to NA do not protect against infection but may attenuate illness (Rott *et al.*, 1974) and antineuraminidase antibodies restrict multiple cycles of replication (Schulman *et al.*, 1968; Webster and Laver, 1967). The active site of the NA is thought to be relatively inaccessible to antibody, since antibody inhibition of enzyme activity has only been observed when large substrates such as fetuin were employed but was not detected with the small substrate, sialyllactose (Rafelson *et al.*, 1963). This implies that the antibody interferes sterically with the approach of the large substrate.

The complete nucleotide sequence of influenza A virus RNA segment 6 which codes for NA has been obtained for both the N1 and N2 subtypes (Fields *et al.*, 1981; Hiti and Nayak, 1982; Markoff and Lai, 1982; Bentley and Brownlee, 1982). Using the PR/8/34 (N1) strain sequence (Fields *et al.*, 1981) as an example (shown schematically in Fig. 6), it has been found that RNA segment 6 is 1413 nucleotides in length with a 5′ non-coding region of 20 nucleotides and a 3′ non-coding region of 31 nucleotides. There is one open reading frame of 1362 nucleotides coding for 453 amino acids, which gives a molecular weight for the NA of 50,087 excluding carbohydrate. There are five possible glycosylation sites (Asn-X-Thr or Asn-X-Ser) and for the N2 strain Tokyo/67 the available data indicate that the carbohydrate units are of both the complex N-acetyl lactosamine and simple oligomannoside type attached via N-glycosidic linkages to asparagine residues (Blok *et al.*, 1982).

The single hydrophobic region in the NA molecule is long enough to span the lipid bilayer of the virus or cell. It is located near the N terminus and consists of 29 residues, of which 18 are hydrophobic, 11 are neutral and none are charged. On

Fig. 6. Schematic diagram of the neuraminidase polypeptide (N1 subtype). The position of the extended signal sequence, the only uncharged domain in the molecule capable of spanning a lipid bilayer is shown. The N-terminal methionine residue that initiates protein synthesis is not removed. Carbohydrate attachment sites are illustrated by the filled circles

the basis of the sequence information alone it was originally suggested that NA was inserted in the membrane by its N terminus (Fields *et al.*, 1981), and this has now been confirmed by protein sequence studies on the intact and protease cleaved "head" molecules (Blok *et al.*, 1982). Thus, NA is similar in orientation to two other glycoproteins, intestinal brush border aminopeptidase and isomaltase (Maroux and Honvard, 1976; Brunner *et al.*, 1979), in having an "extended signal sequence" (Frank *et al.*, 1978; Blobel, 1980) that is not cleaved and which both transfers the protein across the membrane and remains in the bilayer to anchor the protein. The protein sequencing data also indicate that the N-terminal methionine is not removed (Blok *et al.*, 1982), which is an unusual finding.

The nucleotide sequence of RNA segment 6 of influenza B virus, which codes for NA, has also been determined (Shaw *et al.*, 1982) and it shows structural features similar to the NA of the A viruses. An open reading frame starting at the second AUG codon from the 5′ end of the mRNA codes for a protein of 466 amino acids. In most eukaryotic mRNAs, including those of other influenza virus genes sequenced, the first AUG is utilized as the initiation site for translation, although there are cases in which the first AUG codon is not used (reviewed by Kozak, 1981). However, what is unusual in the B/NA sequence is that the first AUG codon, which is separated from the second by four nucleotides, is followed by an open reading frame of 100 amino acids overlapping the NA reading frame (Shaw *et al.*, 1982). A protein product from this reading frame has been recently observed both in infected cells and synthesized *in vitro* (Shaw and Lamb, unpublished observation). The mRNA derived from RNA segment 6 could be bicistronic as both the large protein, NA, and the smaller protein, NB, are synthesized from a mRNA of seemingly identical size. Comparison of the nucleotide and amino acid sequences with those of influenza A NA revealed seven regions of extensive homology within the central position of the NA molecule, including 12 conserved cysteine residues. Five other cysteine residues in the terminal portions were also conserved (Shaw *et al.*, 1982).

The biosynthesis of NA has proved more difficult to examine than that of HA, partly because in many strains of virus the NA comigrates on polyacrylamide gels with the nucleocapsid protein, and it is synthesized in smaller amounts than HA. However, the unglycosylated NA synthesized *in vitro* can be readily detected (Lamb and Choppin, 1979; Lamb and Choppin, 1981). Association of partially glycosylated NA with rough endoplasmic reticulum, and further glycosylation and migration to smooth membranes and the plasma membrane have been observed (Lazarowitz *et al.*, 1971; Compans, 1973 b; Klenk *et al.*, 1974; Hay, 1974). A more detailed analysis, such as described for the VSV G protein (Rothman *et al.*, 1982), should be most interesting in view of the N-terminal membrane insertion of NA.

E. RNA Segment 7: The Structure and Synthesis of the Membrane Protein (M_1) and Non-Structural Protein (M_2)

RNA segment 7 codes for two proteins, M_1 and M_2, and possibly a small peptide (M_3). Three separate mRNAs are derived from RNA segment 7, one of which is a

colinear transcript of the RNA segment and two are spliced mRNAs containing interrupted regions.

The membrane (or matrix) protein (M_1) is the most abundant protein in the virion (Compans *et al.*, 1970; Schulze, 1970; Haslam *et al.*, 1970; Skehel and Schild, 1971; Klenk *et al.*, 1972). Electron microscopy of virions reveals an electron dense layer beneath the lipid bilayer, believed to consist of the M_1 protein (Apostolov and Flewett, 1969; Compans and Dimmock, 1969; Bachi *et al.*, 1969). Further evidence that M_1 underlies the lipid bilayer comes from the observations that proteolytic digestion of virions removes the spikes (HA and NA) but leaves the M_1 protein undigested (Compans *et al.*, 1970) and lipid extraction of fixed virions results in a spikeless shell observable by electron microscopy (Schulze, 1972). Iodination of virions under various conditions indicated that the M_1 protein, while not on the surface, was external to the nucleocapsid (Stanley and Haslam, 1971; Rifkin *et al.*, 1972); this is supported by fluorescent transfer experiments (Lenard *et al.*, 1974). M_1 is the only protein in the virion in sufficient quantity to form a shell beneath the lipid bilayer (Compans *et al.*, 1972; Schulze, 1972). In addition to providing structural stability to the virion envelope, M_1 may recognize the viral glycoproteins and form a domain on the inner surface of the plasma membrane which subsequently provides a binding site for the ribonucleoprotein segments during virus assembly (Choppin *et al.*, 1972). M_1 can interact with lipid, as found by reconstituting lipid and purified M_1 protein (Bucher *et al.*, 1980; Gregoriades, 1980; Gregoriades and Frangione, 1981). M_1 also appears to govern the sensitivity of the virus to the antiviral drug, amantidine (Lubeck *et al.*, 1978; Hay *et al.*, 1979). M_1 is a type-specific antigen for the A viruses and does not cross-react with M, of the B viruses (Schild, 1972; Oxford and Schild, 1976).

The complete sequence of RNA segment 7 of influenza A virus has been obtained for three strains (PR/8/34: Winter and Fields, 1980; Allen *et al.*, 1980; Udorn/72: Lamb and Lai, 1981; FPV/Rostock/34: McCauley *et al.*, 1982). The RNA segment is 1027 nucleotides in length and has a 5′ noncoding region of 25 nucleotides and a 3′ noncoding region of 23 nucleotides. The M_1 protein contains 252 amino acids (MW 27861) and the predicted amino acid sequence fits with the available amino acid composition and partial amino acid sequences (Both and Air, 1979; Robertson *et al.*, 1979). The protein is rich in arginine and methionine and is somewhat hydrophobic, which may account for its solubility in chloroform/

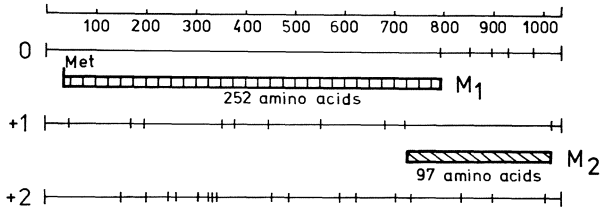


Fig. 7. Schematic diagram of RNA segment 7 of influenza virus. Termination codons in all three reading frames are shown by vertical lines. The two long open reading frames are shown by cross-hatched rectangles. (From Lamb and Lai, 1981)

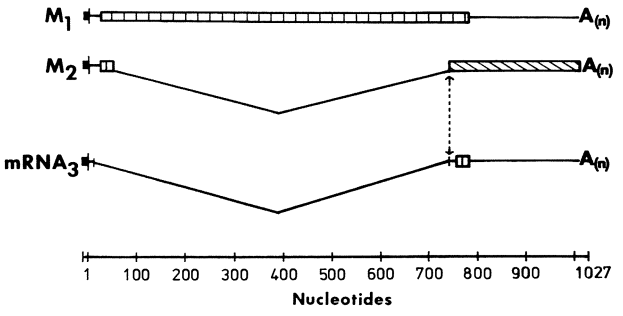


Fig. 8. Model for the arrangement of the M_1, M_2 mRNAs, and M $mRNA_3$ and their coding regions. The thin lines at the 5′ and 3′ termini of the mRNAs represent the non-coding regions. In the region 740–1004, the M_2 mRNA is translated in a reading frame different from that used for M_1. No evidence has yet been obtained that M $mRNA_3$ is translated. The V-shaped lines represent the interrupted regions. The filled-in boxes before nucleotide 1 at the 5′-termini represent the heterogeneous nucleotides derived from cellular mRNAs that are covalently linked to the viral sequences (from Lamb *et al.*, 1981)

methanol (Gregoriades, 1973). There is one region of 37 amino acids in the middle of the molecule possessing 16 hydrophobic amino acids and only two charged amino acids. This region could be involved in hydrophobic interactions with either protein or lipid.

The coding region of the M_1 protein occupies only 75% of the length of RNA segment 7. However, at the 5′ end of the vRNA segment there is a second reading frame which could code for 97 amino acids and which overlaps the first reading frame for the M_1 protein by 68 nucleotides (see Fig. 7) (Winter and Fields, 1980; Allen *et al.*, 1980; Lamb and Lai, 1981). A protein product derived from this reading frame (M_2: apparent MW ~15,000) was identified in influenza virus-infected cells and shown to be encoded in RNA segment 7 by hybrid-arrest translation experiments and by analysis of different recombinants (Lamb and Choppin, 1981; Palese *et al.*, 1981). Polyadenylated mRNAs isolated from influenza virus-infected cells were separated in sucrose gradients and each fraction was translated *in vitro* using a wheat germ extract. Protein M_2 was found to be translated from a distinct small mRNA (Lamb and Choppin, 1981). Nuclease S1 mapping analysis and nucleotide sequencing of the mRNAs showed that the M_2 mRNA contained an interrupted region of 689 nucleotides (Lamb *et al.*, 1981). The 51 virus-specific nucleotides comprising the 5′-terminal leader sequence of the M_2 mRNA are the same as those found at the 5′ end of the colinear M_1 mRNA. Following the leader sequence, there is a 271 nucleotide body region that is 3′-coterminal with the M_1 mRNA. Because the 5′ end sequences of M_1 mRNA and M_2 mRNA are the same and share the 5′-proximal initiation codon for protein synthesis, the first nine amino acids are expected to be the same in the M_1 and M_2 proteins after which the sequences diverge. The 271 nucleotide body region of M_2 mRNA can be translated in the +1 reading frame, and the sequence indicates that M_1 and M_2 overlap by 14 amino acids (Lamb *et al.*, 1981, see Fig. 8). No information about the function of M_2 is available but it is clear that the protein resides only in infected cells and not in virus particles and is

thus a legitimate "non-structural" protein (Lamb and Choppin, 1981). A second interrupted mRNA (M $mRNA_3$) derived from RNA segment 7 has also been found (Lamb *et al.*, 1981; Inglis and Brown, 1981). M $mRNA_3$ has a leader sequence of 11 virus-specific nucleotides that are the same as the 5′ end of the M_1 and M_2 mRNAs, followed by a body region of 271 nucleotides that is the same as that of the M_2 mRNA. The coding potential of the M $mRNA_3$ is for only 9 amino acids, and these would be identical to the C-terminal region of the membrane protein M_1 (see Fig. 8) (Lamb *et al.*, 1981). No evidence for the existence for such a peptide has yet been obtained.

The nucleotides at the 5′ and 3′ junctions of the interrupted M_2 mRNA and M $mRNA_3$ are similar to those of intervening sequences of spliced eukaryotic mRNAs (Seif *et al.*, 1979; Lerner *et al.*, 1980; Lewin, 1980; Mount, 1982) indicating that the interrupted mRNAs are probably produced by splicing of the colinear M_1 mRNA. Evidence that the M_1 mRNA can be spliced is provided by the finding that an identical splice junction is used when mRNAs are transcribed from a cloned DNA copy of RNA segment 7 using a simian virus 40 promoter and cellular enzymes (Lamb and Lai, 1982).

The nucleotide sequence of influenza B virus RNA segment 7 has been determined (Briedis *et al.*, 1982) and is similar in structure to that of the A viruses. RNA segment 7 is 1191 nucleotides in length with an open reading frame of 777 nucleotides coding for the M_1 protein of 248 amino acids. A second coding region overlaps that of the M_1 protein by 86 amino acids in the +2 reading frame and can code for a maximum of 195 amino acids but the size of a protein product depends on whether or not the mRNA contains an interrupted region. It is likely that this region codes for a protein analogous to the influenza A virus M_2 protein, but such a protein has not been identified. Comparison of the amino acid sequences of the M_1 protein between the A and B viruses indicate that 63 amino acids are conserved. The internal hydrophobic domain, which includes three cysteine residues in the identical positions, is also conserved, adding further evidence for its functional significance (Briedis *et al.*, 1982).

The M_1 protein is synthesized on free ribosomes (Compans, 1973 a; Hay, 1974) and as it lacks a signal peptide, its route to the plasma membrane must be different from those taken by HA and NA. Although M_1 has been identified in various subcellular fractions (Lazarowitz *et al.*, 1971; Compans, 1973 a; Klenk *et al.*, 1974; Hay, 1974), there are no indications that it migrates between cellular compartments. Therefore, its recovery from these cell fractions may be artifactual.

F. RNA Segment 8: The Structure and Synthesis of Nonstructural Proteins NS_1 and NS_2

The smallest RNA segment of influenza virus codes for two proteins that are only found in infected cells, nonstructural proteins NS_1 (MW ~ 26,000) and NS_2 (MW ~ 14,000). NS_1 is synthesized in large amounts in infected cells (Lazarowitz *et al.*, 1971; Skehel, 1972) and has been found associated with polysomes and also in the nucleolus (Lazarowitz *et al.*, 1971; Compans, 1973 a; Krug and Etkind, 1973; Krug

and Soeiro, 1975). The function of NS_1 is still unclear but it has been suggested to be involved in the shut off of host cell protein synthesis and virion RNA synthesis (Lazarowitz *et al.*, 1971; Compans, 1973 a; Wolstenholme *et al.*, 1980; Koennecke *et al.*, 1981). NS_1 is a phosphoprotein with phosphate attached to one or two threonine residues per molecule (Privalsky and Penhoet, 1977, 1978, 1981; Almond and Felsenreich, 1982). Late in infection, with some strains of influenza virus, NS_1 forms electron-dense crystalline arrays (inclusions) (Morrongiello and Dales, 1977; Shaw and Compans, 1978), which also contain a mixture of cellular RNA species (Yoshida *et al.*, 1981). These inclusions may have functional consequences, but it seems more likely that they form solely because of the great abundance of NS_1 protein in the dying cell.

A polypeptide designated NS_2 had been observed in infected cells on several occasions (Skehel, 1972; Follett *et al.*, 1974; Minor and Dimmock, 1975; Krug and Etkind, 1973; Stephenson *et al.*, 1977) and was later shown to be a virus-coded unique polypeptide on the basis of its peptide composition and strain-specific differences in its migration in polyacrylamide gels, leading to the suggestion that one virion RNA segment coded for two proteins (Lamb *et al.*, 1978). NS_2 is translated from a separate small mRNA (Lamb and Choppin, 1979; Inglis *et al.*, 1979) and hybrid-arrest translation experiments and analysis of recombinants of defined genome composition showed that NS_2 is encoded by RNA segment 8 (Lamb and Choppin, 1979; Inglis *et al.*, 1979). On the basis of the estimates of the size of RNA segment 8, the sizes of the two distinct proteins NS_1 and NS_2 and from the pattern of translation arrest by DNA restriction fragments of cloned RNA segment 8, it was indicated that NS_1 and NS_2 must be encoded by overlapping reading frames (Lamb and Choppin, 1979; Inglis *et al.*, 1979; Lamb *et al.*, 1980). NS_2 is made late in infection and is located in the cytoplasm of infected cells (Lamb *et al.*, 1978; Lamb and Lai, 1980; Mahy *et al.*, 1980) but its function is unknown.

The complete nucleotide sequence of RNA segment 8 has been determined for several strains (Lamb and Lai, 1980; Porter *et al.*, 1980; Baez *et al.*, 1980, 1981; Winter *et al.*, 1981 b).

Segment 8 is 890 nucleotides in length and, after a 5′ non-coding region of nucleotides, there is an open reading frame coding for the NS_1 protein of 230 to

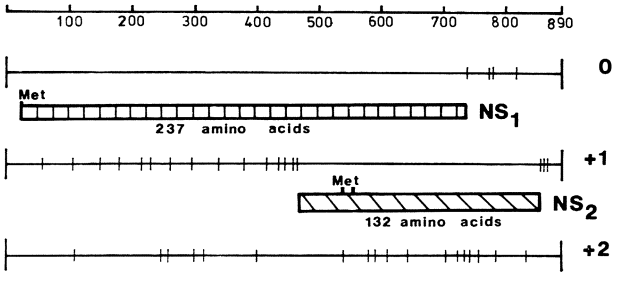


Fig. 9. Schematic representation of RNA segment 8 of influenza virus. Termination codons in all three reading frames are shown by vertical lines. The two long open reading frames are shown by cross-hatched rectangles

237 amino acids depending on the virus strain. The predicted amino acid sequence agrees well with the amino acid composition of NS_1 (Shaw and Compans, 1978) and the tryptic peptides of NS_1 can readily be correlated to the predicted sequence (Lamb *et al.*, 1978; Lamb and Choppin, 1979). The second open reading frame of 132 amino acids found at the 5′ end of virion RNA segment 8 had been predicted from S1 nuclease mapping experiments, hybridization blotting experiments and hybrid arrest translation experiments using cloned DNA of RNA segment 8 (Fig. 9) (Lamb *et al.*, 1980; Inglis *et al.*, 1980). Nucleotide sequencing of NS_2 mRNA showed that it contains an interrupted region of 473 nucleotides. The first 56 virus-specific nucleotides at the 5′ end of the NS_2 mRNA are the same as those found at the 5′ end of NS_1 mRNA. This leader sequence of NS_2 contains the initiation codon for protein synthesis and coding information for nine amino acids which would be common to NS_1 and NS_2. The 340 nucleotide body region of NS_2 mRNA can be translated in the +1 reading frame and the sequence indicates that NS_1 and NS_2 overlap by 70 amino acids that are translated from different reading frames (Fig. 10).

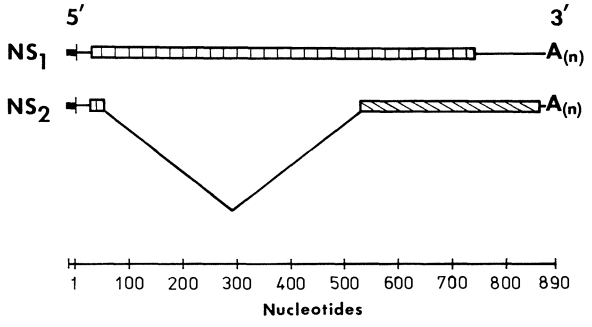


Fig. 10. Schematic representation for the arrangement of the NS_1 and NS_2 mRNAs. The thin lines at the 5′- and 3′-termini of the NS_1 and NS_2 mRNAs represent non-coding regions. The cross-hatched rectangles represent the coding regions of the two mRNAs. In the region 529–861 the NS_2 mRNA is translated in a reading frame different from that used for NS_1. The V-shaped thin line in the NS_2 mRNA represents its interrupted sequence. The short line and bar before nucleotide 1 at the 5′ termini represent the heterogeneous nucleotides derived from cellular mRNAs that are donated to NS_1 and NS_2. (Adapted from Lamb and Lai, 1980)

The nucleotides at the 5′ and 3′ junctions of the interrupted NS_2 mRNA are similar to those of intervening sequences of spliced eukaryotic mRNAs (Lerner *et al.*, 1980; Seif *et al.*, 1979; Lewin, 1980; Mount, 1982), indicating that NS_2 mRNA is probably produced by splicing of the colinear NS_1 mRNA. Evidence for splicing at these nucleotide sequences is provided by finding that the identical splice junction is used when mRNAs are transcribed from cloned DNA of RNA segment 8 using a simian virus 40 promoter and cellular enzymes (Lamb and Lai, in preparation).

The nucleotide sequence of influenza B virus RNA segment 8 has been determined and is 1096 nucleotides in length (Briedis and Lamb, 1982). It also codes for two proteins, NS_1 and NS_2 (Briedis *et al.*, 1981) with an overall structure similar to

that of A viruses. The NS_1 protein has 281 amino acids and NS_2 has the same N terminus of 10 amino acids as NS_1 before the splice junction in the mRNA and then translation continues using the +1 reading frame. The NS_1 protein of influenza B virus is predicted to have a MW of 32,026 but on polyacrylamide gels it migrates anomalously with an apparent MW of 40,000 (Racaniello and Palese, 1979; Briedis *et al.*, 1981). NS_1 contains 122 amino acids, one more than that of influenza A virus, and 52 amino acids overlap between NS_1 and NS_2.

It is interesting to speculate how the overlapping NS_1 and NS_2 genes evolved. One suggestion made by Winter and colleagues (1981 b) is that NS_1 and NS_2 were originally colinear in the vRNA, but not overlapping, Readthrough of a terminator at the end of NS_1 then allowed the NS_1 protein to become longer and use more of the available reading frame. Evidence can be found to support such a concept because (1) the NS_1 protein varies in length depending on the virus strain (2) a mutant virus synthesizes a shortened NS_1 protein (ts 47, Almond *et al.*, 1979), yet seems to replicate normally at the non-permissive temperature and (3) comparison of the nucleotide changes in different strains of virus indicate that in the region of overlap between NS_1 and NS_2, the amino acid sequence of NS_2 is conserved at the expense of NS_1 (Lamb and Lai, 1980; Baez *et al.*, 1980). These data could be interpreted as indicating that the C-terminal region of NS_1 is dispensable to its function.

Examination of the nucleotide sequence of RNA segment 8 of A/PR/8/34, A/Udorn/72 and A/FPV/Rostock/34 showed that the negative (virion sense) strand had an initiation codon at position 98 from the 5′ end and an open reading frame that extended for 167 or 216 amino acids (Baez *et al.*, 1980). It has been speculated that such a reading frame, which statistically should only occur with a very low probability, might be translated (Baez *et al.*, 1980). However, virus RNA segment 8 from neither A/duck/Alberta/60/76 (Baez *et al.*, 1981) nor B/Lee/40 virus (Briedis and Lamb, 1982) contains an open reading frame and therefore it is unlikely to code for a necessary viral protein.

G. Overlapping Coding Regions Using Different Reading Frames in Viruses

The majority of the influenza virus NS_2 polypeptide is translated from a reading frame different from that for NS_1. In the A viruses it overlaps NS_1 by 70 amino acids and in the B viruses it overlaps by 52 amino acids (Lamb and Lai, 1980; Briedis and Lamb, 1982). The majority of the M_2 polypeptide is also translated from a reading frame different from that of M_1 and overlaps M_1 by 14 amino acids (Lamb *et al.*, 1981). Previously, the use of overlapping reading frames has been found with the similar bacteriophages ØX174 and G4 (Barrell *et al.*, 1976; Shaw *et al.*, 1978) in bacteriophage Qβ (Atkins *et al.*, 1979; Beremand and Blumenthal (1979), in the large T and middle T antigens of polyoma virus (Soeda *et al.*, 1979) and in $VP_{2,3}$ and VP_1 of SV40 and polyoma virus (Contreras *et al.*, 1977; Soeda *et al.*, 1980). Whether this efficient use of genomes only occurs with bacteriophages and viruses with small genomes or also occurs in eukaryotic genomes is not yet known.

Acknowledgments

I am very grateful to Andrea Gifford for typing the manuscript, and Mike W. Shaw, Andreas Scheid, and Purnell W. Choppin for helpful comments and encouragement. I thank Peter Palese for providing Fig. 2, Ching-Juh Lai for Fig. 4, and Greg Winter for Fig. 11–18. Research by the author was supported by Research Grants AI-05600 and AI-18746/20201 from the National Institute of Allergy and Infectious Diseases. R. A. L. is an Irma T. Hirschl Awardee and an Established Investigator of the American Heart Association.

Appendix

The Influenza Virus Nucleotide Sequence (A/PR/8/34 Strain)

The complete nucleotide sequence of the 8 RNA segments of influenza A virus (A/PR/8/34) is shown in Fig. 11–18. The original sources of these data are as follows:
RNA segment 1 coding for PB2, Fields and Winter (1982)
RNA segment 2 coding for PB1, Winter and Fields (1982)
RNA segment 3 coding for PA, Fields and Winter (1982)
RNA segment 4 coding for HA, Winter *et al.* (1981 a)
RNA segment 5 coding for NP, Winter and Fields (1981)
RNA segment 6 coding for NA, Fields *et al.* (1981)
RNA segment 7 coding for M_1 and M_2, Winter and Fields (1980)
RNA segment 8 coding for NS_1 and NS_2, Winter *et al.* (1981 b)
The photographs of these sequences, especially formated for this chapter, were kindly made available by Dr. Greg Winter, MRC Laboratory of Molecular Biology, Cambridge, England.
The single-letter amino acid code has been used for the protein sequences:
Alanine (A) Arginine (R) Asparagine (N) Aspartic Acid (D) Cysteine (C) Glutamic acid (E) Glutamine (Q) Glycine (G) Histidine (H) Isoleucine (I) Leucine (L) Lysine (K) Methionine (M) Phenylalanine (F) Proline (P) Serine (S) Threonine (T) Tryptophan (W) Tyrosine (Y) Valine (V)
The nucleotide sequences of the 8 RNA segments are shown in the mRNA sense to facilitate translation into the protein sequence.

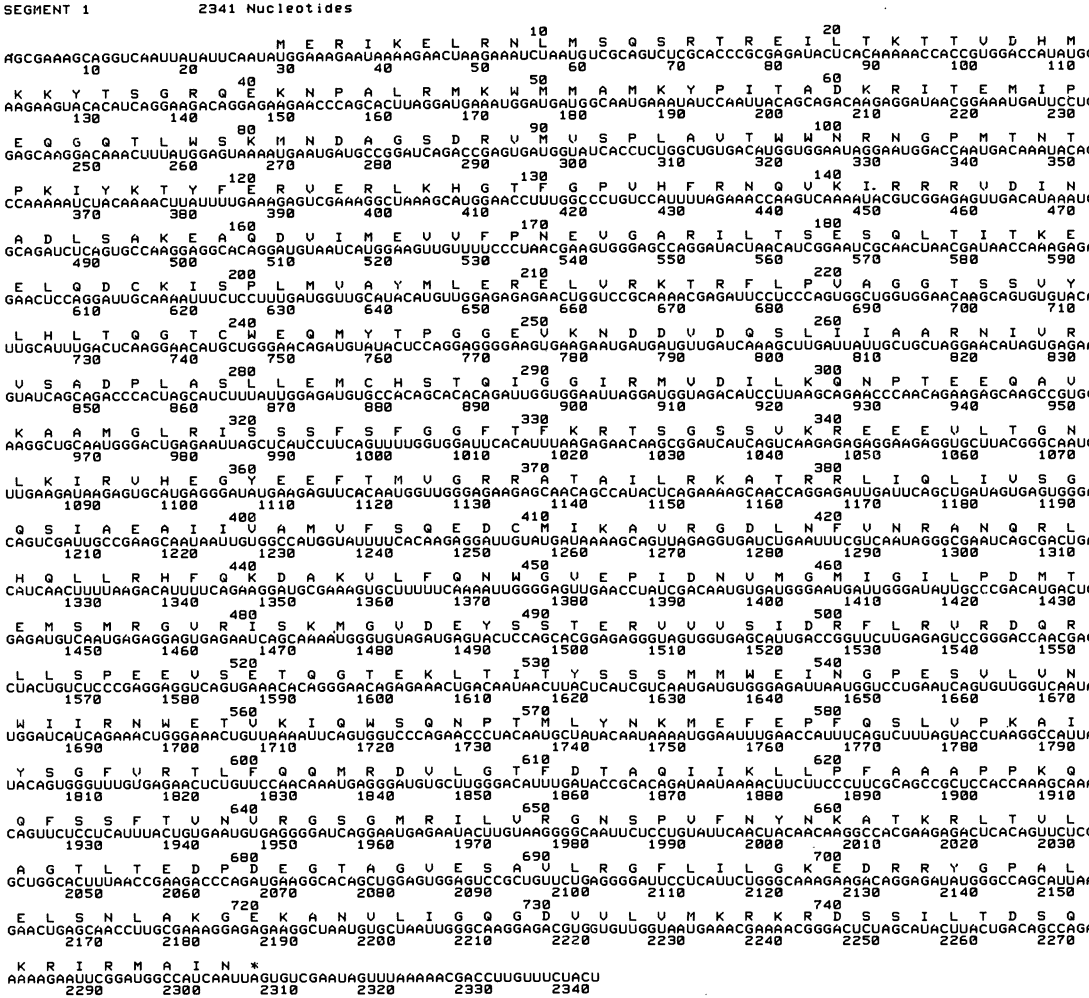


Fig. 11. Sequence of RNA segment 1 coding for PB2. In Figs. 11–18 the nucleotides are shown in the mRNA sense to facilitate translation. The asterisk at the end of the protein sequence indicates the termination of translation codon

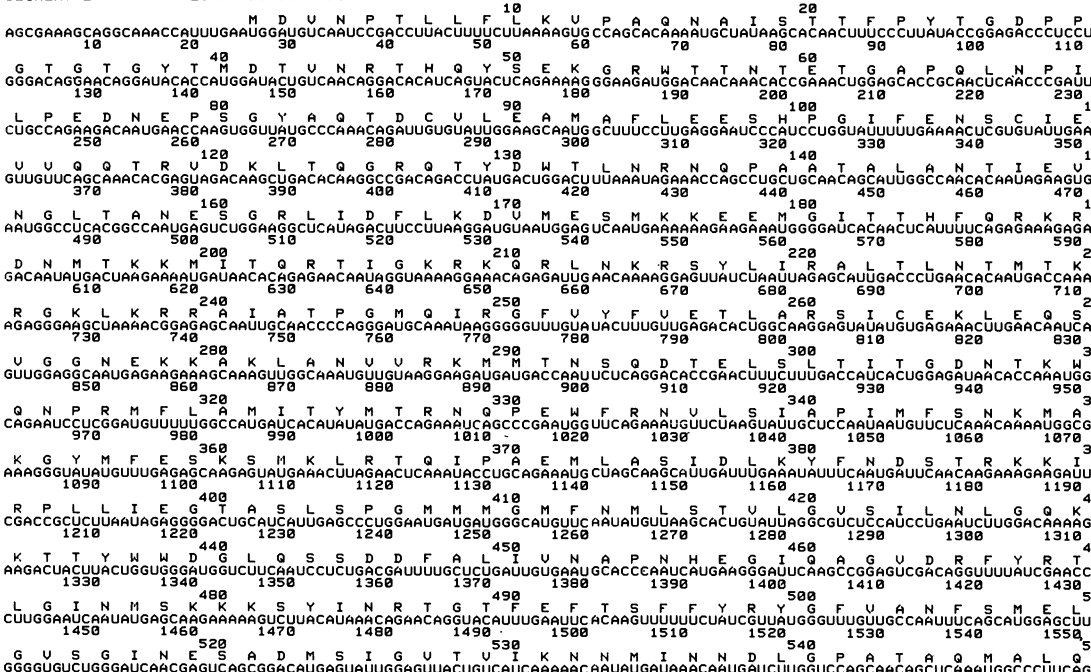


Fig 12. Sequence of RNA segment 2 coding for PB1

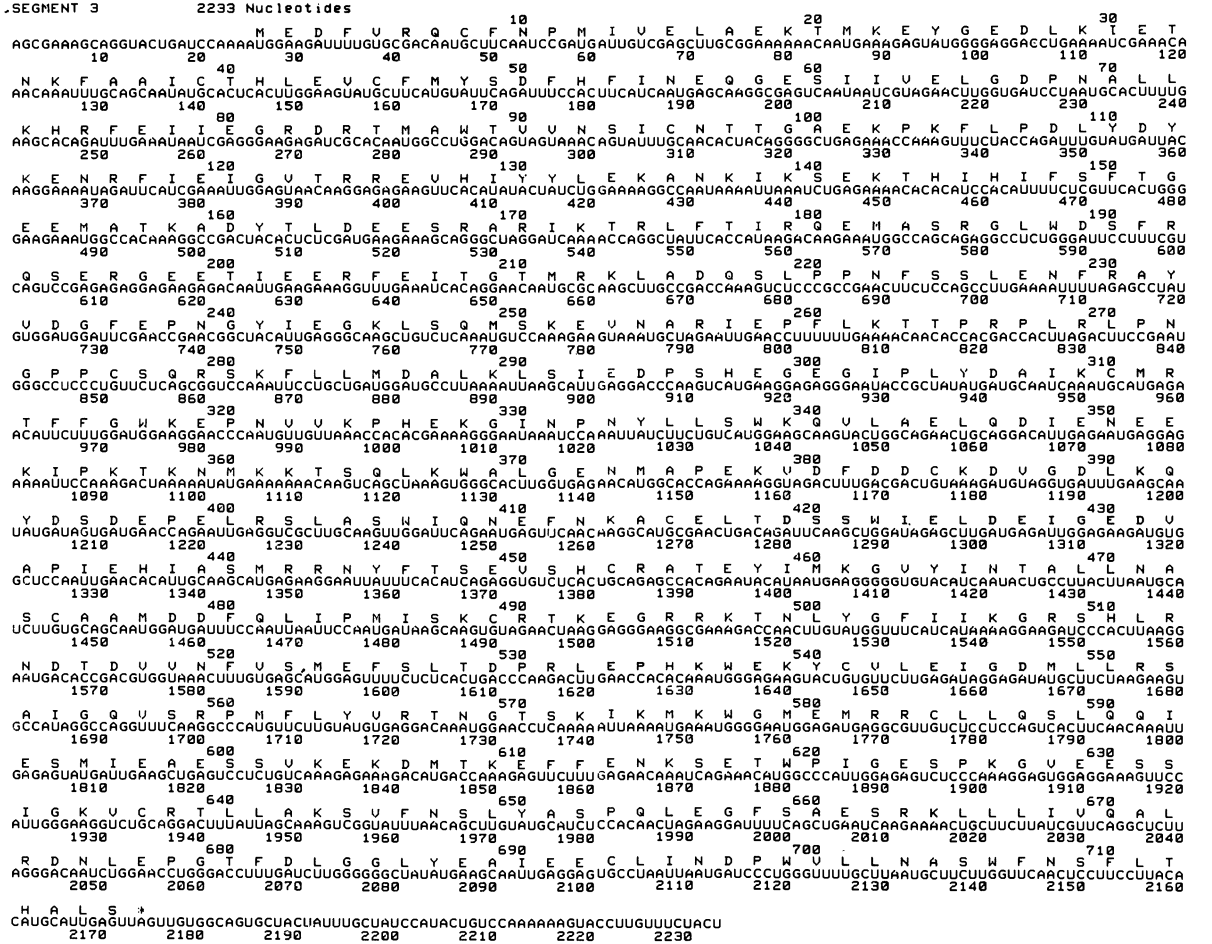


Fig. 13. Sequence of RNA segment 3 coding for PA

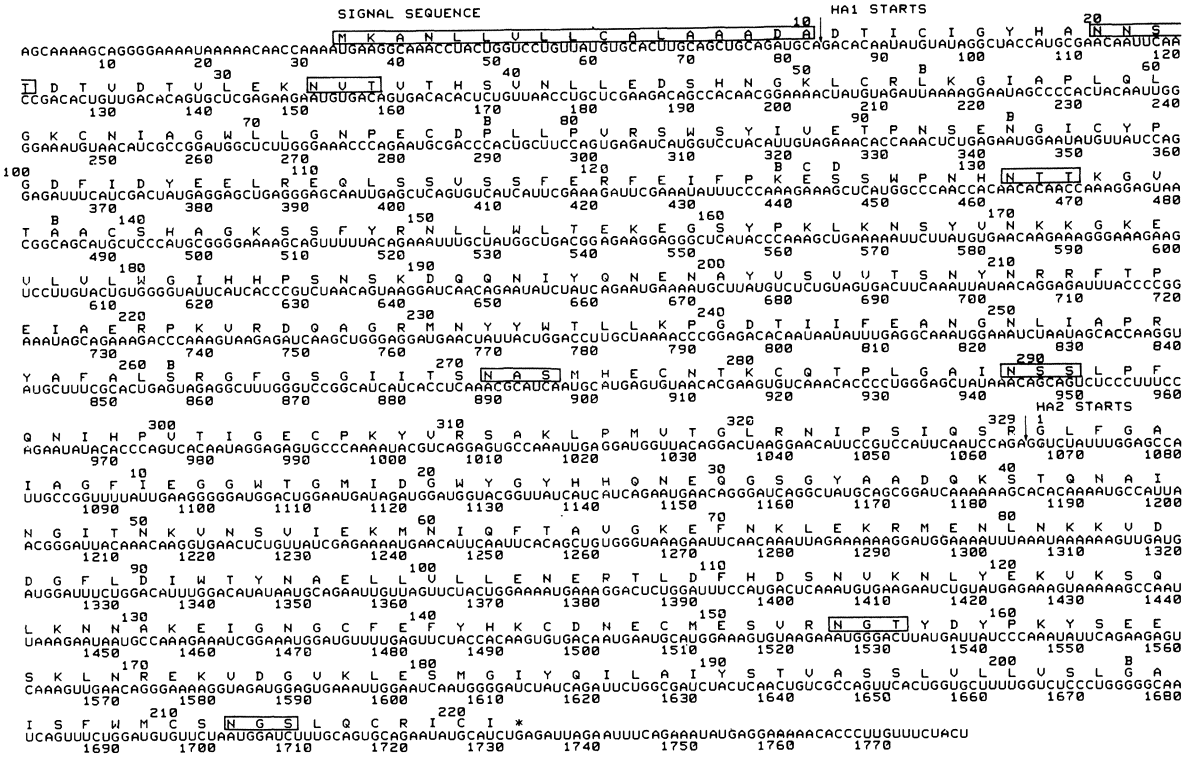


Fig. 14. Sequence of RNA segment 4 coding for HA. The N-terminal signal sequence is shown by the large box. Glycosylation sites are shown by small boxes and the cleavage site of the signal peptide and the site of cleavage of HA into HA1 and HA2 are shown by arrows

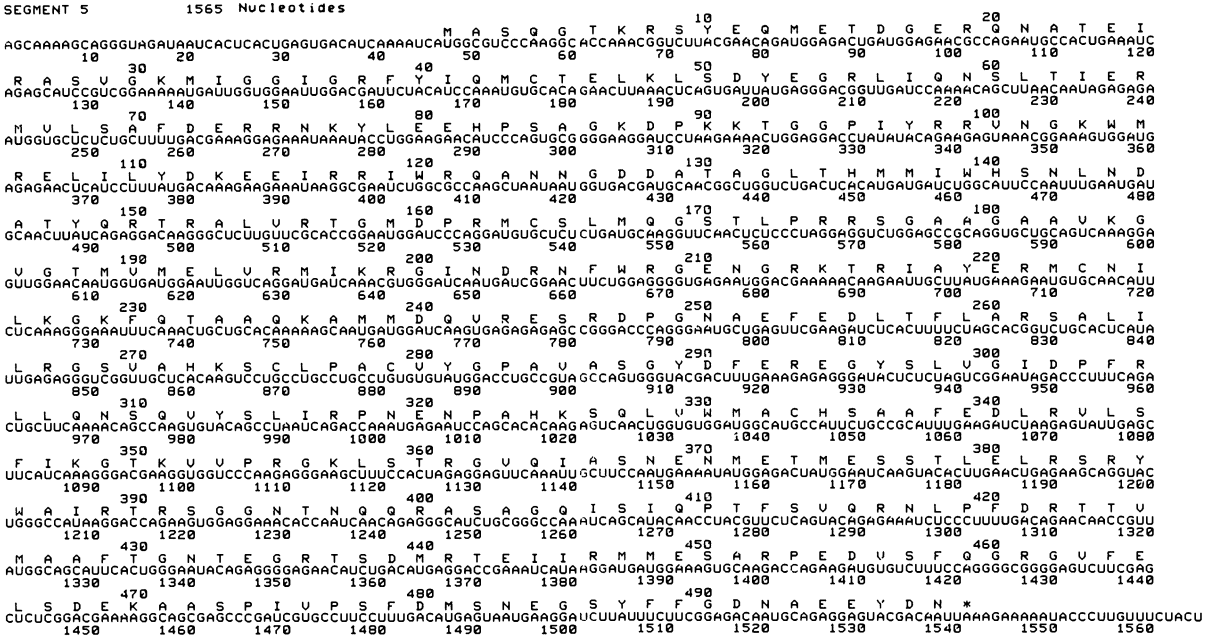


Fig. 15. Sequence of RNA segment 5 coding for NP

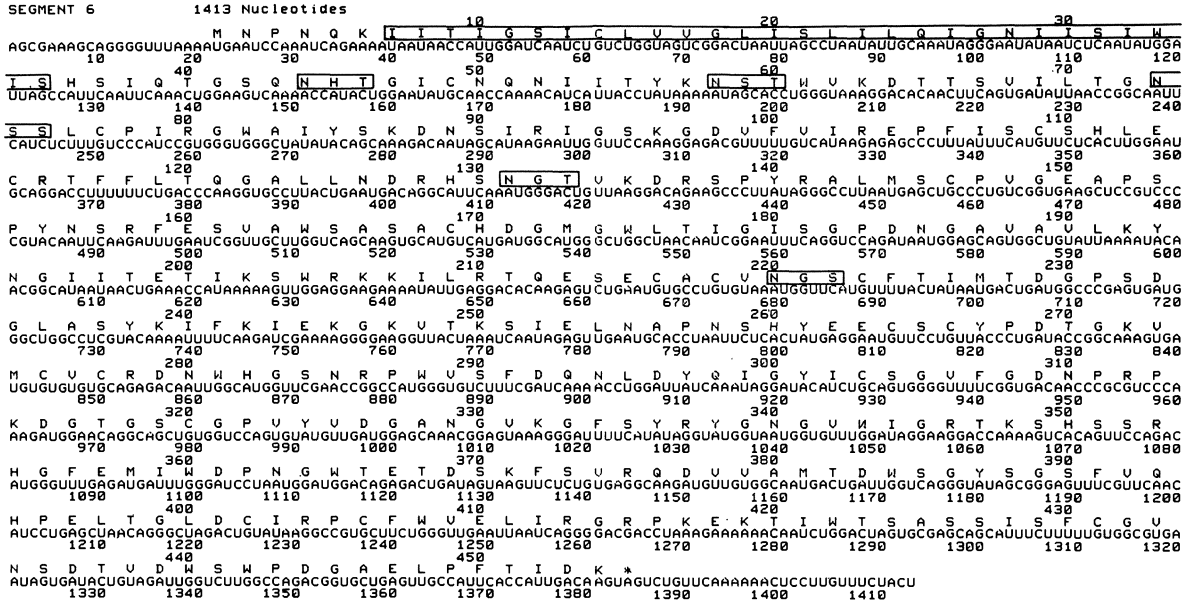


Fig. 16. Sequence of RNA segment 6 coding for NA. The single hydrophobic region in the molecule sufficient to span the lipid-bilayer of the virus or cell is shown by the large box. Glycosylation sites are shown by the small boxes

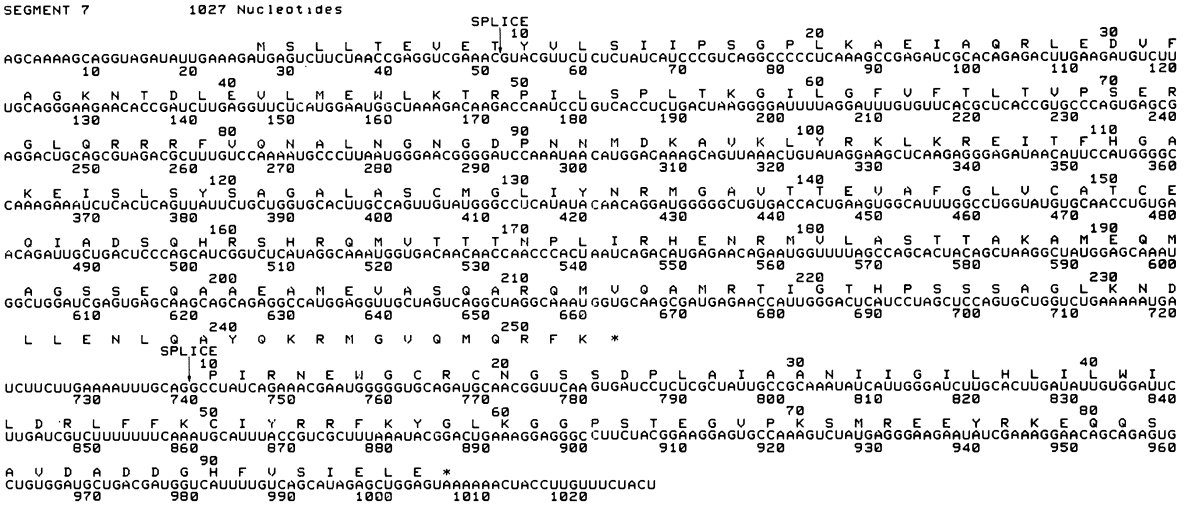


Fig. 17. Sequence of RNA segment 7 coding for M_1 and M_2. The 5′ splice site and 3′ splice site for the M_2 mRNA are shown by arrows. The 5′ splice site in the non-coding region for M $mRNA_3$ is not indicated

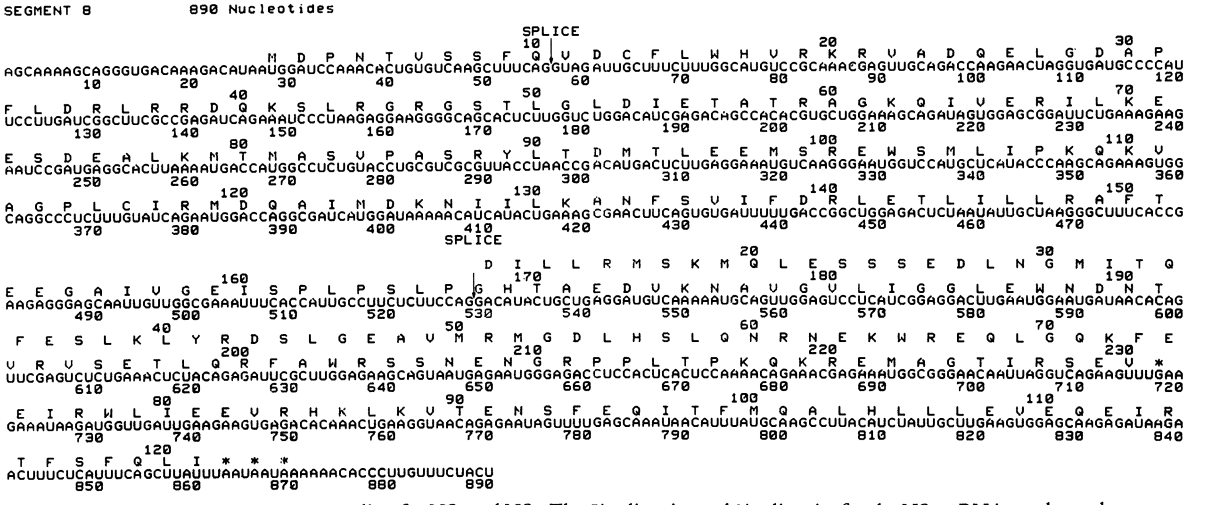


Fig. 18. Sequence of RNA segment 8 coding for NS_1 and NS_2. The 5′ splice site and 3′ splice site for the NS_2 mRNA are shown by arrows

References

Ada, G. L., Perry, B. T.: The nucleic acid content of influenza viruses. Aust. J. Exptl. Biol. Med. Sci. *32,* 453–468 (1954).

Air, G. M: Nucleotide sequence for the signal peptide and N-terminus of the hemagglutinin from an Asian (H2N2) strain of influenza virus. Virology *97,* 468–472 (1979).

Air, G. M.: Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus. Proc. Natl. Acad. Sci. U.S.A. *78,* 7639–7643 (1981).

Air, G. M., Hackett, J. A.: Gene 8 of influenza virus: sequences of cDNA transcribed from the 3′ ends of viral RNA of influenza A and B strains. Virology *103,* 291–298 (1980).

Allen, H., McCauley, J., Waterfield, M., Gething, M. J.: Influenza virus RNA segment 7 has the coding capacity for two polypeptides. Virology *107,* 548–551 (1980).

Almeida, J. D., Brand, C. M.: A morphological study of the internal component of influenza virus. J. Gen. Virol. *27,* 313–318 (1975).

Almond, J. W., Barry, R. D.: Genetic recombination between two strains of fowl plague virus: construction of genetic maps. Virology *92,* 407–415 (1979).

Almond, J. W., Felsenreich, V.: Phosphorylation of the nucleoprotein of an avian influenza virus. J. Gen. Virol. *60,* 295–305 (1982).

Almond, J. W., McGeoch, D., Barry, R. D.: Method for assigning temperature-sensitive mutations of influenza viruses to individual segments of the genome. Virology *81,* 62–73 (1977).

Almond, J. W., McGeoch, D., Barry, R. D.: Temperature-sensitive mutants of fowl plague virus: isolation and genetic characterization. Virology *92,* 416–427 (1979).

Apostolov, K., Flewett, T. H.: Internal structure of influenza virus. Virology *26,* 506–508 (1965).

Atkins, J. F., Steitz, J. A., Anderson, C. W., Model, P.: Binding of mammalian ribosomes to MS2 phage RNA reveals an overlapping gene encoding a lysis function. Cell *18,* 247–256 (1979).

Bachi, T., Gerhard, W., Lindermann, J., Muhlethaler, K.: Morphogenesis of influenza A virus in Erlich ascites tumor cells as revealed by thin sectioning and freeze-etching. J. Virol. *4,* 769–776 (1969).

Baez, M., Zazra, J. J., Elliott, R. M., Young, J. F., Palese, P.: Nucleotide sequence of the influenza A/duck/Alberta/60/76 virus NS RNA: Conservation of the NS1/NS2 overlapping gene structure in a divergent influenza virus RNA segment. Virology *113,* 397–402 (1981).

Baez, M., Taussig, R., Zazra, J. J., Young, J. F., Palese, P., Reisfeld, A., Skalka, A. M.: Complete nucleotide sequence of the influenza A/PR/8/34 virus NS gene and comparison with the NS genes of the A/Udorn/72 and A/FPV Rostock 34 strains. Nucl. Acids Res. *8,* 5845–5858 (1980).

Baltimore, D.: Expression of animal virus genomes. Bacteriol. Rev. *35,* 235–241 (1971).

Barrell, B. G., Air, G. M., Hutchinson, C. A. III: Overlapping genes in bacteriophage $\phi \times 174$. Nature *264,* 34–41 (1976).

Barry, R. D.: The multiplication of influenza virus II. Multiplicity reactivation of UV irradiated virus. Virology 14, 398–405 (1961).

Bean, W. J., jr., Simpson, R. W.: Transcriptase activity and genome composition of defective influenza virus. J. Virol. *18,* 365–369 (1976).

Bentley, D. R., Brownlee, G. G.: Sequence of the N2 neuraminidase from influenza virus A/NT/60/68. Nucl. Acids Res. *10,* 5033–5042 (1982).

Beremand, M. N., Blumenthal, T.: Overlapping genes in RNA phage: a new protein implicated in lysis. Cell *18,* 257–266 (1979).

Bishop, D. H. L., Huddleston, J. A., Brownlee, G. G.: The complete sequence of RNA segment 2 of influenza A/NT/60/68 and its encoded PI protein. Nucl. Acids Res. *10,* 1335–1343 (1982).

Bishop, D. H. L., Obijeski, J. F., Simpson, R. W.: Transcription of the influenza ribonucleic acid genome by a virion polymerase II. Nature of the in vitro polymerase product. J. Virol. *8,* 74–80 (1971).

Bishop, D. H. L., Jones, K. L., Huddleston, J. A., Brownlee, G. G.: Influenza A virus evolution: complete sequences of influenza A/NT/60/68 RNA segment 3 and its predicted acidic P polypeptide compared with those of influenza A/PR/8/34. Virology *120,* 481–489 (1982).

Bishop, D. H. L., Roy, P., Bean, W. J., Simpson, R. W.: Transcription of the influenza ribonucleic acid genome by a virion polymerase III. Completeness of the transcription process. J. Virol. *10,* 689–697 (1972).

Blass, D. B., Patzelt, E., Keuchler, E.: Cap-recognizing protein of influenza virus. Virology *116,* 339–348 (1982).

Blobel, G.: Intracellular protein topogenesis. Proc. Natl. Acad. Sci. U.S.A. *77,* 1496–1500 (1980).

Blok, J., Air, G. M.: Comparative nucleotide sequences at the 3′-end of the neuraminidase gene from eleven influenza type A viruses. Virology *107,* 50–60 (1980).

Blok, J., Air, G. M.: Sequence variation at the 3′ end of the neuraminidase gene from 39 influenza type A viruses. Virology *121,* 211–229 (1982).

Blok, J., Air, G. M., Laver, W. G., Ward, C. W., Lilley, G. G., Woods, E. F., Roxburgh, C. M., Inglis, A. S.: Studies on the size, chemical composition, and partial sequence of the neuraminidase (NA) from type A influenza viruses show that the N-terminal region of the NA is not processed and serves to anchor the NA in the membrane. Virology *119,* 109–121 (1982).

Bosch, F., Orlich, M., Klenk, H.-D., Rott, R.: The structure of the hemagglutinin, a determinant for the pathogenicity of influenza viruses. Virology *95,* 197–207 (1979).

Both, G. W., Air, G. M.: Nucleotide sequence coding for the N-terminal region of the matrix protein of influenza virus. Eur. J. Biochem. *96,* 363–372 (1979).

Both, G. W., Sleigh, M. J.: Complete nucleotide sequence of the haemagglutinin gene from a human influenza virus of the Hong Kong subtype. Nucl. Acids Res. *8,* 2561–2575 (1980).

Both, G. W., Sleigh, M. J.: Conservation and variation in the hemagglutinins of the Hong Kong subtype influenza viruses antigenic drift. J. Virol. *39,* 663–672 (1981).

Brand, C. M., Skehel, J. J.: Crystalline antigen from the influenza virus envelope. Nature (Lond.) New Biol. *238,* 145–147 (1972).

Briedis, D. J., Lamb, R. A.: Influenza B virus genome: sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS_1 and NS_2 proteins. J. Virol. *42,* 186–193 (1982).

Briedis, D. J., Lamb, R. A., Choppin, P. W.: Influenza B virus RNA segment 8 codes for two non-structural proteins. Virology *112,* 417–425 (1981).

Briedis, D. J., Lamb, R. A., Choppin, P. W.: Sequence of RNA segment 7 of the influenza B virus genome: Partial amino acid homology between the membrane proteins (M_1) of influenza A and B viruses and conservation of a second open reading frame. Virology *116,* 581–588 (1982).

Brunner, J., Hauser, H., Braun, H., Wilson, K. J., Wacker, H., O'Neill, B., Semenza, G.: The mode of association of the enzyme complex sucrase isomaltase with the intestinal brush border membrane. J. Biol. Chem. *254,* 1821–1828 (1979).

Bucher, D. J., Kharitonenkow, L. G., Zakomiridin, J. A., Grigoriev, V. B., Klimenko, S. M., Davis, J. F.: Incorporation of influenza M protein into liposomes. J. Virol. *36,* 586–590 (1980).

Burnett, F. M., Lind, P. E.: A genetic approach to influenza virus strains used in mixed infections. J. Gen. Microbiol. *5,* 59–66 (1951).

Burnett, F. M., Lind, P. E.: Genetics of virulence in influenza viruses. Nature *173,* 627–630 (1954).

Caton, A. J., Robertson, J. S.: New procedure for the production of influenza virus-specific double-stranded DNAs. Nucl. Acids Res. *7,* 1445–1456 (1979).

Caton, A. J., Robertson, J. S.: Structure of the host-derived sequences present at the 5′ ends of influenza virus mRNA. Nucl. Acids Res. *8,* 2591–2603 (1980).

Choppin, P. W., Pons, M. W.: The RNAs of infective and incomplete influenza virions grown in MDBK and HeLa cells. Virology *42,* 603–610 (1970).

Choppin, P. W., Murphy, J. S., Stoeckenius, W.: The surface structure of influenza virus filaments. Virology *13,* 549–550 (1961).

Choppin, P. W., Murphy, J. S., Tamm, I.: Studies of two kinds of virus particles which comprise influenza A2 virus strains. III. Morphological characteristics: Independence of morphological and functional traits. J. Exptl. Med. *112,* 945–952 (1960).

Choppin, P. W., Compans, R. W., Scheid, A., McSharry, J. J., Lazarowitz, S. G.: Structure and assembly of viral membranes. In: Membrane Research (Fox, C. F., ed.), 163–179. New York: Academic Press 1972.

Chow, N. L., Simpson, R. W.: RNA dependent RNA polymerase activity associated with virions and subviral particles of myxoviruses. Proc. Natl. Acad. Sci. U.S.A. *68,* 752–756 (1971).

Colman, P. M., Tulloch, P. A., Laver, W. G.: Preliminary structural studies on two influenza virus neuraminidases. In: Structure and Variation in Influenza Viruses (Laver, G., Air, G., eds.), 351–356. New York: Elsevier/North-Holland 1980.

Compans, R. W.: Influenza virus proteins. II. Association with components of the cytoplasm. Virology *51,* 56–70 (1973 a).
Compans, R. W.: Distinct carbohydrate components of influenza virus glycoproteins in smooth and rough cytoplasmic membranes. Virology *55,* 541–545 (1973 b).
Compans, R. W., Caliguiri, L. A.: Isolation and properties of an RNA polymerase from influenza virus-infected cells. J. Virol. *11,* 441–448 (1973).
Compans, R. W., Choppin, P. W.: Reproduction of Myxoviruses. In: Comprehensive Virology (Fraenkel-Conrat, H., Wagner, R. R., eds.), Vol. IV, 179–252. New York: Plenum Press 1975.
Compans, R. W., Dimmock, N. J.: An electron microscope study of single cycle infection of chick embryo fibroblasts by influenza virus. Virology *39,* 499–515 (1960).
Compans, R. W., Bishop, D. H. L., Meier-Ewert, H.: Structural components of influenza C virions. J. Virol. *21,* 658–665 (1977).
Compans, R. W., Content, J., Deusberg, P. H.: Structure of the ribonucleoprotein of influenza virus. J. Virol. *10,* 795–800 (1972).
Compans, R. W., Klenk, H. D., Caliguiri, L. A., Choppin, P. W.: Influenza virus proteins. I. Analysis of polypeptides of the virion and identification of spike glycoproteins. Virology *42,* 880–889 (1970).
Compans, R. W., Roth, M. G., Alonso, F. V.: Directional transport of viral glycoproteins in polarized epithelial cells. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., ed.), 213–231. (ICN-UCLA Symposia on Molecular and Cellular Biology, Vol. XXI.) New York: Academic Press 1981.
Contreras, R., Rogiers, R., Van De Voorde, A., Fiers, W.: Overlapping of the VP_2-VP_3 gene and the VP_1 gene in the SV40 genome. Cell *12,* 529–538 (1977).
Cox, N. J., Kendal, A. P.: Presence of a segmented single-stranded RNA genome in influenza C virus. Virology *74,* 239–241 (1976).
Davies, P., Barry, R. D.: Nucleic acid of influenza virus. Nature *211,* 384–387 (1966).
Desselberger, U., Palese, P.: Molecular weights of RNA segments of influenza A and B viruses. Virology *88,* 394–399 (1978).
Desselberger, U., Racaniello, V. R., Zazra, J. J., Palese, P.: The 3′ and 5′-end terminal sequences of influenza A, B, and C virus RNA segments are highly conserved and show partial inverted complementarity. Gene *8,* 315–328 (1980).
Dhar, R., Chanock, R. M., Lai, C.-J.: Non-viral oligonucleotides at the 5′-terminus of cytoplasmic influenza viral mRNA deduced from cloned complete genomic sequences. Cell *21,* 495–500 (1980).
Dopheide, T. A., Ward, C. W.: The carboxyl-terminal sequence of the heavy chain of a Hong Kong influenza hemagglutinin. Eur. J. Biochem. *85,* 393–398 (1978).
Duesberg, P. H.: The RNAs of influenza virus. Proc. Natl. Acad. Sci. U.S.A. *59,* 930–937 (1968).
Duesberg, P.: Distinct subunits of the ribonucleoprotein of influenza virus. J. Mol. Biol. *42,* 485–499 (1969).
Duesberg, P. H., Robinson, W. S.: On the structure and replication of influenza virus. J. Mol. Biol. *25,* 383–405 (1967).
Elder, K. T., Bye, J. M., Skehel, J. J., Waterfield, M. D., Smith, A. E.: In vitro synthesis, glycosylation and membrane insertion of influenza virus hemagglutinin. Virology *95,* 343–350 (1979).
Emtage, J. S., Catlin, G. H., Carey, N. H.: Polyadenylation and reverse transcription of influenza viral mRNA. Nucl. Acids Res. *6,* 1221–1239 (1979).
Fang, R., Min Jou, W., Huylebroeck, D., Devos, R., Fiers, W.: Complete structure of A/duck/Ukraine/63 influenza hemagglutinin gene: animal virus as progenitor of human H3 Hong Kong 1968 influenza hemagglutinin. Cell *25,* 315–323 (1981).
Fields, S., Winter, G.: Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structure of a small viral RNA segment. Cell *28,* 303–313 (1982).
Fields, S., Winter, G., Brownlee, G. G.: Structure of the neuraminidase gene in human influenza virus A/PR/8/34. Nature *290,* 213–217 (1981).
Flanagan, M. T., Skehel, J. J.: The conformation of influenza virus hemagglutinin. FEBS Lett. *80,* 57–60 (1977).
Floyd, R. W., Stone, M. P., Joklik, W. K.: Separation of single-stranded ribonucleic acids by acrylamide-agarose-urea gel electrophoresis. Analyt. Biochem. *59,* 599–609 (1974).

Follett, E. A. C., Pringle, C. R., Wunner, W. H., Skehel, J. J.: Virus replication in enucleated cells: vesicular stomatitis virus and influenza virus. J. Virol. *13,* 394–399 (1974).

Frank, G., Brunner, J., Hauser, H., Wacker, H., Semenza, G., Zuber, H.: The hydrophobic anchor of small-intestinal sucrase-isomaltase. FEBS Lett. *96,* 183–188 (1978).

Garten, W., Bosch, F. X., Linder, D., Rott, R., Klenk, H.-D.: Proteolytic activation of the influenza virus hemagglutinin: the structure of the cleavage site and the enzymes involved in cleavage. Virology *115,* 361–374 (1981).

Gerhard, W., Yewdell, J., Frankel, M.: An experimental approach to define the antigenic structures of the hemagglutinin molecule of A/PR/8/34. In: Structure and Variation in Influenza Virus (Laver, W. G., Air, G. M., eds.), 273–280. New York: Elsevier/North-Holland 1980.

Gething, M. J., Bye, J., Skehel, J. J., Waterfield, M. D.: Cloning and DNA sequence of double-stranded copies of hemagglutinin genes from H2 and H3 strains elucidates antigenic shift and drift in human influenza virus. Nature *287,* 301–306 (1980).

Gething, M. J., White, J. M., Waterfield, M. D.: Purification of the fusion protein of Sendai virus: analysis of the NH_2-terminal sequence generated during precursor activation. Proc. Natl. Acad. Sci. U.S.A. *75,* 2737–2740 (1978).

Ghendon, Y. Z., Markushin, S. G., Marchenko, A. T., Sitnikov, B. S., Ginzburg, V. P.: Biochemical characterization of fowl plague virus ts mutants. Virology *66,* 454–463 (1973).

Goldstein, E. A., Pons, M. W.: The effect of polyvinylsulfate on the ribonucleoprotein of influenza virus. Virology *41,* 382–384 (1970).

Gottschalk, A.: The specific enzyme of influenza virus and *Vibrio cholerae.* Biochim. Biophys. Acta *23,* 645–646 (1957).

Gregoriades, A.: The membrane protein of influenza virus: extraction from virus and infected cells with acidic chloroform-methanol. Virology *54,* 369–383 (1973).

Gregoriades, A.: Interaction of influenza M protein with viral lipids and phosphatidylcholine vesicles. J. Virol. *36,* 470–479 (1980).

Gregoriades, A., Frangione, B.: Insertion of influenza M protein into the viral lipid bilayer and localization of site of insertion. J. Virol. *40,* 323–328 (1981).

Griffith, I. P.: The fine structure of influenza virus. In: Negative Strand Viruses (Mahy, B. W. J., Barry, R. D., eds.), Vol. 1, 121–132. London: Academic Press 1975.

Gronenborn, B., Messing, J.: Methylation of single-stranded DNA in vitro introduces new restriction endonuclease cleavage sites. Nature *272,* 375–377 (1978).

Haslam, E. A., Hampson, A. W., Radiskevics, I., White, D. O.: The polypeptides of influenza virus. II. Interpretations of polyacrylamide gel electrophoresis patterns. Virology *42,* 555–565 (1970).

Hay, A.: Studies on the formation of the influenza virus envelope. Virology *60,* 398–418 (1974).

Hay, A. J., Kennedy, N. C. T., Skehel, J. J., Appleyard, G.: The matrix protein gene determines amantidine-sensitivity of influenza viruses. J. Gen. Virol. *42,* 189–191 (1979).

Heggeness, M. H., Smith, P. R., Ulmanen, I., Krug, R. M., Choppin, P. W.: Studies on the helical nucleocapsid of influenza virus. Virology *118,* 466–470 (1982).

Hirst, G. K.: Agglutination of red cells by allantoic fluid of chick embryos infected with influenza virus. Science *94,* 22–23 (1941).

Hirst, G. K.: The quantitative determination of influenza virus and antibodies by means of red cell agglutination. J. Exp. Med. *75,* 47–64 (1942).

Hirst, G. K.: The relationship of the receptors of a new strain of virus to those of the mumps-NDV-influenza group. J. Exp. Med. *91,* 177–184 (1950).

Hirst, G. K.: Genetic recombination with Newcastle disease virus, poliovirus and influenza. Cold Spring Harbor Symp. Quant. Biol. *27,* 303–309 (1962).

Hiti, A. L., Nayak, D. P.: Complete nucleotide sequence of the neuraminidase gene of human influenza virus A/WSN/33. J. Virol. *41,* 730–734 (1982).

Hiti, A. L., Davis, A. R., Nayak, D. P.: Complete sequence analysis shows that the hemagglutinins of the H0 and H2 subtypes of human influenza virus are closely related. Virology *3,* 113–124 (1981).

Horisberger, M. A.: The large P proteins of influenza A viruses are composed of one acidic and two basic polypeptides. Virology *107,* 302–305 (1980).

Horne, R. W., Waterson, A. P., Wildy, P., Farnham, A. E.: The structure and composition of the myxo-

viruses. I. Electron microscope studies of the structure of myxovirus particles by negative staining techniques. Virology *11,* 79–98 (1960).
Horst, J., Content, J., Mandeles, S., Frankel-Conrat, H., Deusberg, P.: Distinct oligonucleotide patterns of distinct influenza virus RNA's. J. Mol. Biol. *69,* 209–215 (1972).
Hoyle, L., Horne, R. W., Waterson, A. P.: The structure and composition of the myxoviruses. II. Components released from the influenza virus particle by ether. Virology *13,* 448–459 (1961).
Huang, R. T. C., Rott, R., Klenk, H.-D.: Influenza viruses cause hemolysis and fusion of cells. Virology *110,* 243–247 (1981).
Huang, R. T. C., Wahn, K., Klenk, H.-D., Rott, R.: Fusion between cell membranes and liposomes containing the glycoproteins of influenza virus. Virology *104,* 294–302 (1980).
Huddleston, J. A., Brownlee, G. G.: The sequence of the nucleoprotein gene of human influenza A virus, strain A/NT/60/68. Nucl. Acids Res. *10,* 1029–1038 (1982).
Inglis, S. C., Brown, C. M.: Spliced and unspliced RNAs encoded by virion RNA segment 7 of influenza virus. Nucl. Acids Res. *9,* 2727–2740 (1981).
Inglis, S. C., Gething, M.-J., Brown, C. M.: Relationship between the messenger RNAs transcribed from two overlapping genes of influenza virus. Nucl. Acids Res. *8,* 3575–3589 (1980).
Inglis, S. C., McGeoch, D. J., Mahy, B. W. J.: Polypeptides specified by the influenza virus genome. 2. Assignment of protein coding functions to individual genome segments by *in vitro* translation. Virology *78,* 522–536 (1977).
Inglis, S. C., Barrett, T., Brown, C. M., Almond, J. W.: The smallest genome RNA segment of influenza virus contains two genes that may overlap. Proc. Natl. Acad. Sci. U.S.A. *76,* 3790–3794 (1979).
Inglis, S. C., Carroll, A. R., Lamb, R. A., Mahy, B. W. J.: Polypeptides specified by the influenza virus genome. 1. Evidence for eight distinct gene products specified by fowl plague virus. Virology *74,* 489–503 (1976).
Kamata, T., Wantanabe, Y.: Role for nucleocapsid protein phosphorylation in the transcription of influenza virus genome. Nature *267,* 460–462 (1977).
Kaptein, J. S., Nayak, D. P.: Complete nucleotide sequence of the polymerase 3 gene of human influenza virus A/WSN/33. J. Virol. *42,* 55–63 (1982).
Keil, W., Klenk, H.-D., Schwarz, R. T.: Carbohydrates of influenza virus. III. Nature of oligosaccharide-protein linkage in viral glycoprotein. J. Virol. *31,* 253–256 (1979).
Kendal, A. P.: A comparison of "influenza C" with prototype myxoviruses: receptor-destroying activity (neuraminidase) and structural polypeptides. Virology *65,* 87–99 (1975).
Kilbourne, E. D. (Editor): The influenza viruses and influenza. New York: Academic Press 1975.
Kingsbury, D. W., Webster, R. G.: Some properties of influenza virus nucleocapsids. J. Virol. *4,* 219–225 (1969).
Klenk, H.-D., Choppin, P. W.: Lipids of plasma membranes of monkey and hamster kidney cells and of parainfluenza virions grown in these cells. Virology *38,* 255–268 (1969).
Klenk, H.-D., Choppin, P. W.: Plasma membrane lipids and parainfluenza virus assembly. Virology *40,* 939–947 (1970 a).
Klenk, H.-D., Choppin, P. W.: Glycosphingolipids of plasma membranes of cultured cells and an enveloped virus (SV_5) grown in these cells. Proc. Natl. Acad. Sci. U.S.A. *66,* 57–64 (1970 b).
Klenk, H.-D., Compans, R. W., Choppin, P. W.: An electron microscope study of the presence or absence of neuraminic acid in eveloped viruses. Virology *42,* 1158–1162 (1970).
Klenk, H.-D., Scholtissek, C., Rott, R.: Inhibition of glycoprotein biosynthesis of influenza virus by D-glucosomine and 2-deoxy-glucose. Virology *49,* 723–734 (1972).
Klenk, H.-D., Rott, R., Orlich, M., Biodorn, J.: Activation of influenza A viruses by trypsin treatment. Virology *68,* 426–439 (1975).
Klenk, H.-D., Wollert, W., Rott, R., Scholtissek, C.: Association of influenza virus proteins with cytoplasmic fractions. Virology *57,* 28–41 (1974).
Koennecke, I., Boschek, C. B., Scholtissek, C.: Isolation and properties of a temperature-sensitive mutant (ts412) of an influenza A virus recombinant with a ts lesion in the gene coding for the nonstructural protein. Virology *110,* 16–25 (1981).
Kozak, M.: Mechanism of mRNA recognition by eukaryotic ribosomes during initiation of protein synthesis. In: Current Topics in Microbiology and Immunology, Vol. 93, 81–123. Berlin-Heidelberg-New York: Springer 1981.

Krug, R. M., Etkind, P. R.: Cytoplasmic and nuclear specific proteins in influenza virus-infected MDCK cells. Virology *56,* 334–348 (1973).

Krug, R. M., Soeiro, R.: Studies on the intranuclear localization of influenza virus-specific proteins. Virology *64,* 378–387 (1975).

Krug, R. M., Ueda, M., Palese, P.: Temperature-sensitive mutants of influenza WSN virus defective in virus-specific RNA synthesis. J. Virol. *16,* 790–796 (1975).

Krystal, M., Elliott, R. M., Benz, E. W., jr., Young, J. F., Palese, P.: Evolution of influenza A and B viruses: Conservation of structural features in the hemagglutinin genes. Proc. Natl. Acad. Sci. U.S.A. *79,* 4800–4804 (1982).

Lai, C.-J., Markoff, L. J., Zimmerman, S., Cohen, B., Berndt, J. A., Chanock, R. M.: Cloning DNA sequences from influenza viral RNA segments. Proc. Natl. Acad. Sci. U.S.A. *77,* 210–214 (1980).

Lamb, R. A., Choppin, P. W.: Synthesis of influenza virus proteins in infected cells: Translation of viral polypeptides, including three P polypeptides, from RNA produced by primary transcription. Virology *74,* 504–519 (1976).

Lamb, R. A., Choppin, P. W.: Segment 8 of the influenza virus genome is unique in coding for two polypeptides. Proc. Natl. Acad. Sci. U.S.A. *76,* 4908–4912 (1979).

Lamb, R. A., Choppin, P. W.: Identification of a second protein (M_2) encoded by RNA segment 7 of influenza virus. Virology *112,* 729–737 (1981).

Lamb, R. A., Lai, C.-J.: Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. Cell *21,* 475–485 (1980).

Lamb, R. A., Lai, C.-J.: 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 (1981).

Lamb, R. A., Lai, C.-J.: Spliced and unspliced messenger RNAs synthesized from cloned influenza virus M DNA in an SV40 vector: expression of the influenza virus membrane protein (M_1). Virology *123,* 237–256 (1982).

Lamb, R. A., Etkind, P. R., Choppin, P. W.: Evidence for a ninth influenza viral polypeptide. Virology *91,* 60–78 (1978).

Lamb, R. A., Lai, C.-J., Choppin, P. W.: Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. Proc. Natl. Acad. Sci. U.S.A. *78,* 4170–4174 (1981).

Lamb, R. A., Choppin, P. W., Chanock, R. M., Lai, C.-J.: Mapping of the two overlapping genes for polypeptides NS_1 and NS_2 on RNA segment 8 of influenza virus genome. Proc. Natl. Acad. Sci. U.S.A. *77,* 1857–1861 (1980).

Laver, W. G.: Structural studies on the protein subunits from three strains of influenza virus. J. Mol. Biol. *9,* 109–124 (1964).

Laver, W. G.: Separation of two polypeptide chains from the hemagglutinin subunit of influenza virus. Virology *45,* 275–288 (1971).

Laver, W. G., Kilbourne, E. D.: Identification in a recombinant virus of structural proteins derived from both parents. Virology *30,* 493–501 (1966).

Laver, W. G., Valentine, R. C.: Morphology of the isolated hemagglutinin and neuraminidase subunits of influenza virus. Virology *38,* 105–119 (1969).

Laver, W. G., Air, G. M., Webster, R. G.: The mechanisms of antigenic drift in influenza virus. Amino acid sequence changes in an antigenically active region of Hong Kong (H3N2) influenza virus hemagglutinin. J. Mol. Biol. *145,* 339–361 (1981).

Laver, W. G., Air, G. M., Dopheide, T. A., Ward, C. W.: Amino acid sequence changes in the haemagglutinin of A/Hong Kong (H3N2) influenza virus during the period 1968–1977. Nature *283,* 454–457 (1980).

Laver, W. G., Gerhard, W., Webster, R. G., Frankel, M. E., Air, G: Antigenic drift in type A influenza virus: peptide mapping and antigenic analysis of A/PR/8/34 (HONI) variants selected with monoclonal antibodies. Proc. Natl. Acad. Sci. U.S.A. *76,* 1425–1429 (1979).

Lazarowitz, S. G., Choppin, P. W.: Enhancement of infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. Virology *68,* 440–454 (1975).

Lazarowitz, S. G., Compans, R. W., Choppin, P. W.: Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes. Virology *46,* 830–843 (1971).

Lazarowitz, S. G., Compans, R. W., Choppin, P. W.: Proteolytic cleavage of the hemagglutinin polypeptide of influenza virus: function of the uncleaved polypeptide HA. Virology *52,* 199–212 (1973 a).

Lazarowitz, S. G., Goldberg, A. R., Choppin, P. W.: Proteolytic cleavage of plasmin of the HA polypeptide of influenza virus. Host cell activation of serum plasminogen. Virology *56,* 172–180 (1973 b).

Lazdins, I., Haslam, E. A., White, D. O.: The polypeptides of influenza virus. VI. Composition of the neuraminidase. Virology *49,* 758–765 (1972).

Lenard, J., Miller, D. K.: pH dependent hemolysis by influenza, Semliki forest and Sendai virus. Virology *110,* 479–482 (1981).

Lenard, J. C., Wong, C. Y., Compans, R. W.: Association of the internal membrane protein with the lipid bilayer in influenza virus. A study with the fluorescent probe 12-(9-anthroyl)-stearic acid. Biochim. Biophys. Acta *332,* 341–349 (1974).

Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L., Steitz, J. A.: Are snRNPs involved in splicing? Nature *283,* 220–224 (1980).

Lewandowski, L. J., Content, J., Leppla, S. H.: Characterization of the subunit structure of the ribonucleic acid genome of influenza virus. J. Virol. *8,* 701–710 (1971).

Lewin, B.: Alternatives for splicing: recognizing the ends of introns. Cell *22,* 324–326 (1980).

Lin, B.-C., Lai, C.-J.: The influenza virus nucleoprotein synthesized from cloned DNA in an SV40 vector is detected in the nucleus. J. Virol. *45,* 434–438 (1982).

Lubeck, M. D., Schulman, J. L., Palese, P.: Susceptibility of influenza A viruses to amantadine is influenced by the gene coding for M protein. J. Virol. *28,* 710–716 (1978).

Mackenzie, J. S.: Isolation of temperature-sensitive mutants and the construction of a preliminary genetic map of influenza virus. J. Gen. Virol. *6,* 63–75 (1970).

Maeda, T., Kawasaki, K., Ohnishi, S.-I.: Interaction of influenza virus hemagglutinin with target membrane lipids is a key step in virus-induced hemolysis and fusion at pH 5.2. Proc. Natl. Acad. Sci. U.S.A. *78,* 4133–4137 (1981).

Mahy, B. W. J., Barrett, T., Briedis, D. J., Brownson, J. M., Wolstenholme, A. J.: Influence of the host cell on influenza virus replication. Phil. Trans. R. Soc. (Lond.) *B 288,* 349–357 (1980).

Markoff, L., Lai, C.-J.: Sequence of the influenza A/Udorn/72 (H3N2) virus neuraminidase gene as determined from cloned full-length DNA. Virology *119,* 288–297 (1982).

Maroux, S., Louvard, D.: On the hydrophobic part of aminopeptidase and maltases which bind the enzyme to the intestinal brush border membrane. Biochim. Biophys. Acta *419,* 189–195 (1976).

Matlin, K., Reggio, H., Helenius, A., Simons, K.: Infectious entry pathway of influenza virus in a canine kidney cell line. J. Cell Biol. *91,* 601–613 (1981).

Mayron, L. W., Robert, B., Winzler, R. J., Rafelson, M. E.: Studies on the neuraminidase of influenza virus. 1. Separation and some properties of the enzyme from Asian and PR8 strains. Arch. Biochem. Biophys. *92,* 475–483 (1961).

Maxam, A. M., Gilbert, W.: A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. *74,* 560–564 (1977).

McCauley, J. W., Mahy, B. W. J., Inglis, S. C.: Nucleotide sequence of fowl plaque virus RNA segment 7. J. Gen. Virol. *58,* 211–215 (1982).

McCauley, J., Skehel, J. J., Elder, K., Gething, M.-J., Smith, A., Waterfield, M.: Haemagglutinin biosynthesis. In: Structure and Variation in Influenza Viruses (Laver, G., Air, G., eds.), 97–104. New York: Elsevier/North-Holland 1980.

McCauley, J., Bye, J., Elder, K., Gething, M. J., Skehel, J. J., Smith, A., Waterfield, M. D.: Influenza virus hemagglutinin signal sequences. FEBS Lett. *108,* 422–428 (1979).

McClelland, L., Hare, R.: The adsorption of influenza virus by red cells and a new in vitro method of measuring antibodies for influenza virus. Can. J. Public Health *32,* 530–538 (1941).

McGeoch, D. J., Fellner, P., Newton, C.: The influenza virus genome consists of eight distinct RNA species. Proc. Natl. Acad. Sci. U.S.A. *73,* 3045–3049 (1976).

McMaster, G. K., Carmichael, G. G.: Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. Proc. Natl. Acad. Sci. U.S.A. *74,* 4835–4838 (1977).

Meier-Ewert, H., Compans, R. W., Bishop, D. H. L., Herrler, G.: Molecular analyses of influenza C

viruses. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 127–133. London: Academic Press 1978.

Mills, J. V., Chanock, R. M.: Temperature-sensitive mutants of influenza virus. 1. Behavior in tissue culture and in experimental animals. J. Infect. Dis. *123,* 145–157 (1971).

Min Jou, W., Verhoeyen, M., Devos, R., Saman, E., Fang, R., Huylebroeck, D., Fiers, W., Threlfall, G., Barber, C., Carey, N., Emtage, S.: Complete structure of the hemagglutinin gene from the human influenza A/Victoria/3/75 (H3N2) strain as determined from cloned DNA. Cell *19,* 683–696 (1980).

Minor, P. D., Dimmock, N. J.: Inhibition of synthesis of influenza virus proteins: evidence for two host-cell dependent events during multiplication. Virology *67,* 114–123 (1975).

Morgan, C., Hsu, K. C., Rifkind, R. A., Knox, A. W., Rose, H. M.: The application of ferritin-conjugated antibody to electron microscopic studies of influenza virus in infected cells. 1. The cellular surface. J. Exp. Med. *114,* 825–837 (1961).

Morrongiello, M. P., Dales, S.: Characterization of cytoplasmic inclusions formed during influenza/WSN virus infection of chick embryo fibroblasts. Intervirology *8,* 281–293 (1977).

Mount, S. M.: A catologue of splice junction sequences. Nucl. Acids Res. *10,* 459–472 (1982).

Murphy, J. S., Bang, F. B.: Observations with the electron microscope on cells of the chick chorio-allantoic membrane infected with influenza virus. J. Exp. Med. *95,* 259–271 (1952).

Murti, K. G., Bean, W. J., jr., Webster, R. G.: Helical ribonucleoproteins of influenza virus: an electron microscope analysis. Virology *104,* 224–229 (1980).

Nakajima, S., Sugiura, A.: Neurovirulence of influenza virus in mice. II. Mechanism of virulence as studied in a neuroblastoma cell line. Virology *101,* 450–457 (1980).

Nakamura, K., Compans, R. W.: Glycopeptide components of influenza viral glycoproteins. Virology *86,* 432–442 (1978).

Nakamura, K., Compans, R. W.: Host cell and virus strain-dependent differences in oligo-saccharides of hemagglutinin glycoproteins of influenza A viruses. Virology *95,* 8–23 (1979).

Nakamura, K., Herrler, G., Petri, T., Meier-Ewert, H., Compans, R. W.: Carbohydrate components of influenza C virions. J. Virol. *29,* 997–1005 (1979).

Nerome, K., Ishida, M.: The multiplication of an influenza C virus in an established line of canine kidney (MDCK) cells. J. Gen. Virol. *39,* 179–181 (1978).

Noll, H., Aoyagi, T., Orlando, J.: The structural relationship of sialidase to the influenza virus surface. Virology *18,* 154–157 (1962).

Oxford, J. S., Schild, G. C.: Immunological and physicochemical studies of influenza matrix (M) polypeptides. Virology *74,* 394–402 (1976).

Palese, P., Schulman, J. L.: Isolation and characterization of influenza virus recombinants with high and low neuraminidase activity: Use of 2-(3′-methoxyphenyl)-N-acetylneuraminic acid to identify cloned populations. Virology *57,* 227–237 (1974).

Palese, P., Schulman, J. L.: Differences in RNA patterns of influenza A viruses. J. Virol. *17,* 876–884 (1976 a).

Palese, P., Schulman, J. L.: Mapping of the influenza virus genome: identification of the hemagglutinin and the neuraminidase genes. Proc. Natl. Acad. Sci. U.S.A. *73,* 2142–2146 (1976 b).

Palese, P., Ritchey, M. B., Schulman, J. L.: Mapping of the influenza virus genome. II. Identification of the P1, P2, and P3 genes. Virology *76,* 114–121 (1977 a).

Palese, P., Ritchey, M. B., Schulman, J. L.: P1 and P3 proteins of influenza virus are required for complementary RNA synthesis. J. Virol. *21,* 1187–1195 (1977 b).

Palese, P., Tohita, K., Ueda, M., Compans, R. W.: Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. Virology *61,* 397–410 (1974).

Palese, P., Elliott, R. M., Baez, M., Zazra, J. J., Young, J. F.: Genome diversity among influenza A, B, and C viruses and genetic structure of RNA 7 and RNA 8 of influenza A viruses. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., ed.) (ICN-UCLA Symposia on Molecular and Cellular Biology, Vol. XXI), 127–140. New York: Academic Press 1981.

Penn, C. R., Blaas, D., Kuechler, E., Mahy, B. W. J: Identification of the cap-binding protein of two strains of influenza A/FPV. J. Gen. Virol. *62,* 177–180 (1982).

Petri, T., Dimmock, N. J.: Phosphorylation of influenza virus nucleoprotein in vivo. J. Gen. Virol. *57,* 185–190 (1981).

Plotch, S. J., Bouloy, M., Ulmanen, I., Krug, R. M.: Initiation of influenza viral RNA transcription by capped RNA primers: a unique cap (m^7GpppXm)-dependent virion endonuclease generates 5′ terminal RNA fragments that prime transcription. Cell *23,* 847–858 (1981).

Pons, M. W.: Isolation of influenza virus ribonucleoprotein from infected cells. Demonstration of the presence of negative stranded RNA in viral RNP. Virology *46,* 149–160 (1971).

Pons, M. W.: A re-examination of influenza single- and double-stranded RNAs by gel electrophoresis. Virology *69,* 789–792 (1976).

Pons, M. W., Hirst, G. K.: Polyacrylamide gel electrophoresis of influenza virus RNA. Virology *34,* 385–388 (1968).

Pons, M. W., Schulze, I. T., Hirst, G. K.: Isolation and characterization of the ribonucleoprotein of influenza virus. Virology *39,* 250–259 (1969).

Porter, A. G., Smith, J. C., Emtage, J. S.: Nucleotide sequence of influenza virus RNA segment 8 indicates that coding regions for NS_1 and NS_2 proteins overlap. Proc. Natl. Acad. Sci. U.S.A. *77,* 5074–5078 (1980).

Porter, A. G., Barber, C., Carey, N. H., Hallewell, R. A., Threfall, G., Emtage, J. S.: Complete nucleotide sequence of an influenza virus hemagglutinin gene from cloned DNA. Nature *282,* 471–477 (1979).

Privalsky, M. L., Penhoet, E. E.: Phosphorylated protein component present in influenza virions. J. Virol. *24,* 401–405 (1977).

Privalsky, M. L., Penhoet, E. E.: Influenza virus proteins: identitly, synthesis, and modification analyzed by two-dimensional gel electrophoresis. Proc. Natl. Acad. Sci. U.S.A. *75,* 3625–3629 (1978).

Privalsky, M. L., Penhoet, E. E.: The structure and synthesis of influenza virus phosphoproteins. J. Biol. Chem. *256,* 5368–5376 (1981).

Rancaniello, V. R., Palese, P.: Influenza B virus genome: assignment of viral polypeptides to RNA segments. J. Virol. *29, 29,* 361–373 (1979).

Rafelson, M. E., Schneir, M., Wilson, V. M.: Studies on the neuraminidase of influenza virus. Arch. Biochem. Biophys. *103,* 424–430 (1963).

Rees, P. J., Dimmock, N. J.: Electrophoretic separation of influenza virus ribonucleoproteins. J. Gen. Virol. *53,* 125–132 (1981).

Richardson, C. D., Scheid, A., Choppin, P. W.: Specific inhibition of paramyxovirus and myxovirus replication by oligopeptides with amino acid sequences similar to those at the N-terminal of the F_1 or HA_2 viral polypeptides. Virology *105,* 205–222 (1980).

Rifkin, D. B., Compans, R. W., Reich, E.: A specific labeling procedure for proteins on the outer surface of membranes. J. Biol. Chem. *247,* 6432–6437 (1972).

Ritchey, M. B., Palese, P., Kilbourne, E. D.: The RNAs of influenza A, B, and C viruses. J. Virol. *18,* 738–744 (1976 a).

Ritchey, M. B., Palese, P., Schulman, J. L.: Mapping of the influenza virus genome. III. Identification of genes coding for nucleoprotein, membrane protein, and nonstructural protein. J. Virol. *20,* 307–313 (1976 b).

Ritchey, M. B., Palese, P., Schulman, J. L.: Differences in protein patterns of influenza A viruses. Virology *76,* 122–128 (1977).

Robertson, J. S.: 5′ and 3′ terminal nucleotide sequences of the RNA genome segments of influenza virus. Nucl. Acids Res. *6,* 3745–3757 (1979).

Robertson, J. S., Schubert, M., Lazzarini, R. A.: Polyadenylation sites for influenza mRNA. J. Virol. *38,* 157–163 (1981).

Robertson, B. H., Bhown, A. S., Compans, R. W., Bennett, J. C.: Structure of the membrane protein of influenza virus. 1. Isolation and characterization of cyanogen bromide cleavage products. J. Virol. *30,* 759–766 (1979).

Rochovansky, O. M.: RNA synthesis by ribonucleoprotein-polymerase complexes isolated from influenza virus. Virology *73,* 327–338 (1976).

Rohde, W., Harms, E., Scholtissek, C.: Biochemical studies on influenza viruses. 1. Comparative analysis of equine 2 virus N genes and gene products. Virology *79,* 393–404 (1977).

Rodriguez-Boulan, E., Sabatini, D. D.: Asymmetric budding of viruses in epithelial monolayers: a model system for study of epithelial polarity. Proc. Natl. Acad. Sci. U.S.A. *75,* 5071–5075 (1978).

Rohde, W., Harms, E., Scholtissek, C.: Studies on the genome structure of influenza A viruses. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 11–17. London: Academic Press 1978.

Roth, M. G., Fitzpatrick, J. P., Compans, R. W.: Polarity of influenza and vesicular stomatitis virus motivation in MDCK cells. Lack of requirement for glycosylation of viral glycoproteins. Proc. Natl. Acad. Sci. *76,* 6430 (1979).

Rothman, J. E., Fries, E., Dunphy, W. G., Urbani, L. J.: The Golgi apparatus, coated vesicles and the sorting problem. Cold Spring Harbor Symposia on Quantitive Biology, Volume XLVI: Organization of the cytoplasm, 797–805 (1982).

Rott, R.: Molecular basis of infectivity and pathogenicity of myxovirus. Arch. Virol. *59,* 285–298 (1979).

Rott, R., Klenk, H.-D.: Structure and assembly of viral envelopes. In: Virus infection and the cell surface. Cell surface reviews (Poste, G., Nicholson, G. L., eds.), Vol. 2, 47–81. Amsterdam: Elsevier/North-Holland 1977.

Rott, R., Becht, H., Orlich, M.: The significance of influenza virus neuraminidase in immunity. J. Gen. Virol. *22,* 35–41 (1974).

Rott, R., Reinacher, M., Orlich, M., Klenk, H.-D.: Cleavability of hemagglutinin determines spread of avian influenza viruses in the chorioallantoic membrane of chicken embryo. Arch. Virol. *65,* 123–133 (1980).

Sanger, F., Brownlee, G., Barrell, B. G.: A two-dimensional fractionation procedure for radioactive nucleotides. J. Mol. Biol. *13,* 373–398 (1965).

Sanger, F., Nicklen, S., Coulson, A. R.: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. *74,* 5463–5467 (1977).

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., Roe, B. A.: Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. *143,* 161–178 (1980).

Scheid, A., Choppin, P. W.: Identification of the biological activities of paramyxovirus glycoproteins. Activation of cell fusion, hemolysis, and infectivity by proteolytic cleavage of an inactive precursor protein of Sendai virus. Virology *57,* 475–490 (1974).

Scheid, A., Choppin, P. W.: Two disulfide-linked polypeptide chains constitute the active F protein of paramyxoviruses. Virology *80,* 54–66 (1977).

Scheid, A., Graves, M., Silver, S. M., Choppin, P. W.: Studies on the structure and function of paramyxovirus glycoproteins. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 181–193. London: Academic Press 1978.

Schild, G. C.: Evidence for a new type-specific structural antigen of the influenza virus particle. J. Gen. Virol. *15,* 99–103 (1972).

Schild, G. C., Oxford, J. S., Newman, R. W.: Evidence for antigenic variation in influenza A nucleoprotein. Virology *93,* 569–573 (1979).

Scholtissek, C., Bowles, A. L.: Isolation and characterization of temperature-sensitive mutants of fowl plague virus. Virology *67,* 576–587 (1975).

Scholtissek, C., Rott, R.: Behavior of virus-specific activities in tissue cultures infected with myxoviruses after chemical changes of the viral ribonucleic acid. Virology *22,* 169–179 (1964).

Scholtissek, C., Kruczinna, R., Rott, R., Klenk, H.-D.: Characteristics of an influenza mutant temperature-sensitive for viral RNA synthesis. Virology *58,* 317–322 (1974).

Scholtissek, C., Harms, E., Rohde, W., Orlich, M., Rott, R.: Correlation between RNA fragments of fowl plague virus and their corresponding gene functions. Virology *74,* 332–344 (1976).

Schulman, J. L., Palese, P.: Selection and identification of influenza virus recombinants of defined genetic composition. J. Virol. *20,* 248–254 (1976).

Schulman, J. L., Palese, P. J.: Virulence factors of influenza A viruses: WSN virus neuraminidase required for productive infection in MDBK cells. J. Virol. *24,* 170–176 (1977).

Schulman, J. L., Khakpour, M., Kilbourne, E. D.: Protective effects of specific immunity to viral neuraminidase on influenza virus infection of mice. J. Virol. *2,* 778–786 (1968).

Schulze, I. T.: The structure of influenza virus. 1. The polypeptides of the virion. Virology *42,* 890–904 (1970).

Schulze, I. T.: The structure of influenza virus. II. A model based on the morphology and composition of subviral particle. Virology *47,* 181–196 (1972).

Schwarz, R. T., Schmidt, M. F. G., Anwer, U., Klenk, H.-D.: Carbohydrates of influenza virus. I. Glycopeptides derived from viral glycoproteins after labelling with radioactive sugars. J. Virol. *23,* 217–226 (1977).

Seif, I., Khoury, G., Dhar, R.: BKV splice sequences based on analysis of preferred donor and acceptor sites. Nucl. Acids Res. *6,* 3387–3398 (1979).

Shaw, M. W., Compans, R. W.: Isolation and characterization of cytoplasmic inclusions from influenza A virus-infected cells. J. Virol. *25,* 605–615 (1978).

Shaw, M. W., Lamb, R. A., Erickson, B. W., Briedis, D. J., Choppin, P. W.: Complete nucleotide sequence of the neuraminidase gene of influenza B virus. Proc. Natl. Acad. Sci. U.S.A. *79,* 6817–6871 (1982).

Shaw, D. C., Walker, J. E., Northrop, F. D., Barrell, B. C., Godson, G. N., Fiddes, J. C.: Gene K, a new overlapping gene in bacteriophage G4. Nature *272,* 510–515 (1978).

Simpson, R. W., Hirst, G. K.: Genetic recombination among influenza viruses. 1. Cross reactivation of plaque-forming capacity as a method for selecting recombinants from the progeny of crosses between influenza A strains. Virology *15,* 436–451 (1961).

Simpson, R. W., Hirst, G. K.: Temperature-sensitive mutants of influenza A virus: Isolation of mutants and preliminary observations on genetic recombination and complementation. Virology *35,* 41–49 (1968).

Sippel, A. E.: Purification and characterization of adenosine triphosphate: ribonucleic acid adenyltransferase from Escherichia coli. Eur. J. Biochem. *37,* 31–40 (1973).

Sivasubramanian, N., Nayak, D. P.: Sequence analysis of the polymerase 1 gene and the secondary structure prediction of polymerase 1 protein of human influenza virus A/WSN/33. J. Virol. *44,* 321–329 (1982).

Skehel, J. J.: Polypeptide synthesis in influenza virus-infected cells. Virology *49,* 23–36 (1972).

Skehel, J. J., Hay, A. J.: Nucleotide sequences of the 5′ termini of influenza virus RNAs and their transcripts. Nucl. Acids Res. *5,* 1207–1219 (1978).

Skehel, J. J., Schild, G. C.: The polypeptide composition of influenza A viruses. Virology *44,* 396–408 (1971).

Skehel, J. J., Waterfield, M. D.: Studies on the primary structure of the influenza virus hemagglutinin. Proc. Natl. Acad. Sci. U.S.A. *72,* 93–97 (1975).

Skehel, J. J., Bayley, P. M., Brown, E. B., Martin, S. R., Waterfield, M. D., White, J. M., Wilson, I. A., Wiley, D. C.: Changes in the confirmation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. Proc. Natl. Acad. Sci. U.S.A. *79,* 968–972 (1982).

Sleigh, M. J., Both, G. W., Brownlee, G. G.: A new method for the size estimation of the RNA genome segments of influenza virus. Nucl. Acids Res. *6,* 1309–1321 (1979).

Sleigh, M. J., Both, G. W., Brownlee, G. G., Bender, V. J., Moss, B. A.: The haemagglutinin gene of influenza A virus: nucleotide sequence analysis of cloned DNA copies. In: Structure and Variation in Influenza Viruses (Laver, G., Air, G. M., eds.), 69–79. New York: Elsevier/North-Holland 1980.

Smith, J. C., Carey, N. H., Fellner, P., McGeoch, D., Barry, R. D.: Comparative studies of nucleotide sequences within two influenza virus genes. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 37–46. London: Academic Press 1978.

Soeda, E., Arrand, J. R., Griffin, B. E.: Polyoma virus DNA: complete nucleotide sequence of the gene which codes for polyoma virus capsid protein VP1 and overlaps the VP2/VP3 genes. J. Virol. *33,* 619–630 (1980).

Soeda, E., Arrand, J. R., Smolar, N., Griffin, B. E.: Sequence from early regions of polyoma virus DNA containing viral replication origin and encoding small, middle, and (part of) large T antigens. Cell *17,* 357–370 (1979).

Stanley, P. M., Haslam, E. A.: The polypeptides of influenza virus. V. Localization of polypeptides in the virion by iodination techniques. Virology *46,* 764–773 (1971).

Stephenson, J. R., Hay, A. J., Skehel, J. J.: Characterization of virus-specific messenger RNAs from avian fibroblasts infected with fowl plague virus. J. Gen. Virol. *36,* 237–248 (1977).

Sugiura, A., Ueda, M.: Neurovirulence of influenza viruses in mice. 1. Neurovirulence of recombinants between virulent and avirulent virus strains. Virology *101,* 440–449 (1980).

Sugiura, A., Tobita, K., Kilbourne, E. D.: Isolation and preliminary characterization of temperature-sensitive mutants of influenza virus. J. Virol. *10,* 639–647 (1972).

Sveda, M. M., Markoff, L. J., Lai, C.-J.: Cell surface expression of the influenza virus hemagglutinin requires the hydrophobic carboxy-terminal sequences. Cell *30,* 649–656 (1982).

Taniguchi, T., Palmieri, M., Weissmann, C.: Qβ DNA-containing hybrid plasmids giving rise to Qβ phage formation in the bacterial host. Nature *274,* 223–228 (1978).

Ueda, M.: Temperature-sensitive mutants of influenza virus. Isolation and preliminary characterization. Arch. ges. Virusforsch. *39,* 360–368 (1972).

Ulmanen, I., Broni, B. A., Krug, R. M.: Role of two of the influenza virus core P proteins in recognizing cap 1 structures (m^7GpppNm) on RNAs and in initiating viral RNA transcription. Proc. Natl. Acad. Sci. U.S.A. *78,* 7355–7359 (1981).

Ulmanen, I., Broni, B., Krug, R. M.: Influenza virus temperature-sensitive cap (m^7GpppNm)-dependent endonuclease. J. Virol. *45,* 27–35 (1983).

Van Rompuy, L., Min Jou, W., Huylebroeck, D., Devos, R., Fiers, W.: Complete nucleotide sequence of the nucleoprotein gene from the human influenza strain A/PR/8/34 (HONI). Eur. J. Biochem. *116,* 347–353 (1981).

Van Wyke, K. L., Hinshaw, V. S., Bean, W. J., Webster, R. G.: Antigenic variation of influenza A virus nucleoprotein detected with monoclonal antibodies. J. Virol. *35,* 24–30 (1980).

Varghese, J. N., Laver, W. G., Colman, P. M.: Structure of the influenza virus glycoprotein antigen neuraminidase at 2–9 Å resolution. Nature *303,* 35–40 (1983).

Verhoeyen, M., Fang, R., Min Jou, W., Devos, R., Huylebroeck, D., Saman, E., Fiers, W.: Antigenic drift between the haemagglutinin of the Hong Kong influenza strains A/Aichi/2/68 and A/Victoria/3/75. Nature *286,* 771–776 (1980).

Ward, C. W.: Structure of the influenza virus hemagglutinin. In: Current Topics in Microbiology and Immunology, Vol. 95/95, 1–74. Berlin-Heidelberg-New York: Springer 1981.

Ward, C. W., Dopheide, T. A.: Primary structure of the Hong Kong (H3) hemagglutinin. Brit. Med. Bull. *35,* 51–56 (1979).

Ward, C. W., Dopheide, T. A. A.: Amino acid sequence and oligosaccharide distribution of the hemagglutinin from an early Hong Kong variant A/Aichi/2/68 (X-31). Biochem. J. *193,* 953–962 (1981).

Waterfield, M. D., Espelie, K., Elder, K., Skehel, J. J.: Structure of the haemagglutinin of influenza virus. Brit. Med. Bull. *35,* 57–63 (1979).

Waterfield, M. D., Gething, M.-J., Scrace, G., Skehel, J. J.: The carbohydrate side chains and disulfide bonds of the hemagglutinin of the influenza virus A/Japan/305/57 (H2N1). In: Structure and Variation in Influenza Virus (Laver, G., Air, G. M., eds.), 11–20. New York: Elsevier/North-Holland 1980.

Webster, R. G., Laver, W. G.: Preparation and properties of antibody directed specifically against the neuraminidase of influenza virus. J. Immunol. *99,* 49–55 (1967).

Webster, R. G., Laver, W. G.: Determination of the number of non-overlapping antigenic areas of Hong Kong (H3N2) influenza virus hemagglutinin with monoclonal antibodies and the selection of variants with potential epidemiological significance. Virology *104,* 139–148 (1980).

Webster, R. G., Laver, W. G., Air, G. M., Schild, G. C.: Molecular mechanisms of variation in influenza viruses. Nature *296,* 115–121 (1982).

White, J., Kartenbeck, J., Helenius, A.: Membrane fusion activity of influenza virus. EMBO Journal *1,* 217–222 (1982).

Wiley, D. C., Skehel, J. J., Waterfield, M. D.: Evidence from studies with a cross-linking reagent that the hemagglutinin of influenza virus is a trimer. Virology *79,* 446–448 (1977).

Wiley, D. C., Wilson, I. A., Skehel, J. J.: Structural identification of the antibody binding sites of the Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature *298,* 373–378 (1981).

Wilson, V. W., Rafelson, M. E.: Isolation of neuraminidase from influenza virus. Biochem. Prep. *10,* 113–117 (1963).

Wilson, I. A., Skehel, J. J., Wiley, D. C.: The hemagglutinin membrane glycoprotein of influenza virus: structure at 3 Å resolution. Nature *289,* 366–373 (1981).

Winter, G., Fields, S.: Cloning of influenza cDNA into M13: the sequence of the RNA segment encoding the A/PR/8/34 matrix protein. Nucl. Acids Res. *8,* 1965–1974 (1980).

Winter, G., Fields, S.: The structure of the gene encoding the nucleoprotein of human influenza virus A/PR/8/34. Virology *114,* 423–428 (1981).

Winter, G., Fields, S.: Nucleotide sequence of human influenza A/PR/8/34 segment 2. Nucl. Acids Res. *10,* 2135–2143 (1982).

Winter, G., Fields, S., Brownlee, G. G.: Nucleotide sequence of the haemagglutinin of a human influenza virus H1 subtype. Nature *292,* 72–75 (1981 a).

Winter, G., Fields, S., Gait, M. J., Brownlee, G. G.: The use of synthetic oligodeoxynucleotide primers in cloning and sequencing segment 8 of influenza virus (A/PR/8/34). Nucl. Acids Res. *9,* 237–245 (1981 b).

Wolstenholme, A. J., Barrett, T., Nichol, S. T., Mahy, B. W. J.: Influenza virus-specific RNA and protein synthesis in cells infected with temperature-sensitive mutants defective in the genome segment encoding nonstructural proteins. J. Virol. *35,* 1–7 (1980).

Wrigley, N. G., Laver, W. G., Downie, J. C.: Binding of antibodies to isolated hemagglutinin and neuraminidase molecules of influenza virus observed in the electron microscope. J. Mol. Biol. *109,* 405–421 (1977).

Wrigley, N. G., Skehel, J. J.., Charlwood, P. A., Brand, C. M.: The size and shape of influenza virus neuraminidase. Virology *51,* 525–529 (1973).

Yewdell, J. W., Webster, R. G., Gerhard, W.: Antigenic variation in three distinct determinants of an influenza type A hemagglutinin molecule. Nature *279,* 246–248 (1979).

Yoshida, T., Shaw, M., Young, J. F., Compans, R. W.: Characterization of the RNA associated with influenza A cytoplasmic inclusions and the interaction of NS_1 protein with RNA. Virology *110,* 87–97 (1981).

Young, R. J., Content, J.: 5′-terminus of influenza virus RNA. Nature New Biol. *230,* 140–142 (1971).

Young, J. F., Desselberger, U., Graves, P., Palese, P., Shatzman, A., Rosenberg, M.: Cloning and expression of influenza virus genes. In: The Origin of Pandemic Viruses (Laver, W. G., ed.). New York: Elsevier/North-Holland 1983.

3

Transcription and Replication of Influenza Viruses

R. M. Krug

I. Introduction

The segmented RNA genome of influenza virus is of negative polarity, *i.e.*, the viral messenger RNA (mRNA) is complementary to the genome or virion RNA (vRNA) and the virion contains the enzyme system which transcribes the vRNA into the viral mRNA [75]. The synthesis of influenza viral mRNA involves a unique interaction with the host cell transcriptional machinery in the nucleus of the infected cell. This interaction is required first for the initiation of the synthesis of the viral mRNA chains. A viral endonuclease cleaves 5′-terminal fragments from newly synthesized capped (m^7GpppNm-containing) cellular RNAs in the nucleus. These are most likely heterogeneous nuclear RNAs (hnRNAs), the precursors of cellular mRNAs [33, 42, 77]. These fragments of capped host nuclear RNAs serve as primers to initiate viral mRNA synthesis. The interaction with host cell nuclear functions apparently continues after the viral mRNAs are synthesized. The viral mRNAs, like other mRNAs (both viral and cellular) synthesized in the nucleus [81], contain internal N6 methyl adenosine (m^6A) residues [43, 46], and several of the viral mRNAs appear to be generated by splicing like that occurring during the processing of hnRNAs to form cellular mRNAs [50]. Most likely, internal methylation and splicing of influenza viral mRNAs are carried out by cellular RNA processing enzymes in the nucleus.

Because the viral mRNAs contain host-derived sequences at their 5′ ends and lack sequences complementary to the last 17 to 22 nucleotides at the 5′ ends of the vRNA segments [29, 31, 42], the viral mRNAs are not suitable templates for the synthesis of vRNA (*i.e.*, replication). The presumed templates for vRNA replication are full-length transcripts of the vRNA segments. These full-length transcripts, which lack 3′-terminal polyadenylate (poly A) sequences, comprise only 5% of the viral transcripts synthesized in the infected cell [29]. The synthesis of the fulllength transcripts requires the synthesis of one or more viruscoded proteins. The initiation of the full-length transcripts apparently occurs without a primer [32]. Also, the termination of transcription that occurs during viral mRNA synthesis 17 to 22 nucleotides from the 5′ end of vRNA must be prevented.

Very little is known about vRNA synthesis that is directed by the full-length transcripts.

This chapter will review what is presently known about the synthesis of the three types of virus-specific RNAs—viral mRNA, full-length transcripts and vRNA—and will point out many of the remaining unanswered questions.

II. Viral mRNA Synthesis

A. Priming by Cellular Capped RNAs—Discovery

For many years, evidence accumulated indicating the influenza virus is unique among nononcogenic RNA viruses in requiring the functioning of host nuclear RNA polymerase II, the enzyme which synthesizes the precursors to cellular mRNAs [5, 52, 64, 85, 95]. The most definitive evidence was that α-amanitin, a specific inhibitor of RNA polymerase II, inhibits virus replication and that in mutant cells containing an α-amanitin-resistant RNA polymerase II, virus replication is also resistant to this drug [52, 85, 95]. The activity of RNA polymerase II was shown to be required for viral RNA transcription, even primary transcription: when added at the beginning of infection, α-amanitin inhibits all detectable viral RNA transcription [67].

An explanation for the RNA polymerase II requirement was proposed on the basis of studies of the transcription reaction catalyzed *in vitro* by the virion-associated transcriptase. It was found that the virion transcriptase is unable to initiate RNA synthesis effectively without the addition of a primer. In initial studies, specific dinucleoside monophosphates, ApG or GpG, at relatively high concentrations (about 0.2 mM) were found to act as primers to initiate chains [68, 78]. Transcription is strongly stimulated (as much as 100-fold) by ApG or GpG [68, 78], and the resulting viral RNA transcripts contain poly (A) and function as viral mRNAs in cell-free systems [11, 78]. With these dinucleotides, transcription initiates exactly at the UC sequence at the 3′ ends of the vRNA segments [80, 93]. It was also found that the virion-associated transcriptase complex lacks enzymes capable of capping the 5′ terminus of the viral RNA transcripts even when these transcripts contain di- or triphosphorylated 5′ ends supplied by a ppApG or a pppApG primer, respectively [80]. However, it was known that, like most eukaryotic mRNAs [3, 91], viral mRNAs isolated from the infected cell do contain 5′-terminal methylated cap structures [46]. Based on these results, it was proposed that viral mRNA synthesis *in vivo* also requires a primer, but that the *in vivo* primer is not a dinucleotide but rather an RNA synthesized by RNA polymerase II, and that the 5′ cap is derived from this primer RNA [78–80].

Direct evidence in support of this hypothesis was first obtained using the virion-associated transcriptase system. Initially, it was shown that the addition of purified β-globin mRNA to the virion-associated transcriptase stimulates viral RNA transcription *in vitro* about 80-fold [11]. On a molar basis, β-globin mRNA was found to be about 1000 times more effective as a primer than ApG. Subsequently, it was found that all RNAs tested that contain 5′-terminal methylated cap structures

were effective primers for viral mRNA synthesis [19, 11, 76]. The presence of the 5′-terminal methylated cap structure (cap 1, $m^7GpppNm$) is required for priming activity [76]. The cap must contain methyl groups, since 5′-GpppG-terminated RNAs are inactive as primers [10]. In fact, each of the methyl groups in the cap, the 2′-O-methyl on the penultimate base as well as the 7-methyl on the terminal G, strongly influence priming activity [12]. As predicted by the original hypothesis, the cap of the primer RNA is transferred to the viral mRNA during transcription [10, 76]. To demonstrate this, a globin mRNA was prepared that contained ^{32}P only in its cap [76]. After *in vitro* transcription in the presence of this primer RNA and unlabeled nucleoside triphosphates, the resulting viral mRNAs were shown to contain ^{32}P-labeled cap structures. In addition about 10 to 15 of the 5′-terminal nucleotides of the primer RNA were also transferred to the viral mRNAs. Initially, this was demonstrated in two ways: (i) gel electrophoretic analysis indicated that globin mRNA-primed viral mRNAs are 10 to 15 nucleotides larger than ApG-primed viral mRNAs [76]; and (ii) using ^{125}I-labeled β-globin mRNA as primer, it was shown that the viral mRNAs contain at their 5′ ends either the first 12, 13, or 14 5′-terminal nucleotides of β-globin mRNA [82].

Subsequently, it was established that viral RNA transcription in the infected cell is also primed by capped cellular RNAs. Several lines of evidence indicated that the viral mRNAs synthesized in the infected cell contain a short stretch of nucleotides at their 5′ ends, including the cap, that are not viral-coded. First, gel electrophoresis showed that the *in vivo* viral mRNAs are 10 to 15 nucleotides longer at their 5′ ends that the ApG-primed *in vitro* viral mRNAs [43]. Second, when ^3H-methyl-labeled *in vivo* viral mRNA was hybridized to vRNA, the 5′-terminal cap structure of the mRNA was not protected against release from the hybrids by pancreatic or T1 ribonuclease (RNase) digestion [43]. Finally, cloning experiments directly demonstrated the existence of nonviral nucleotides at the 5′ ends of the *in vivo* viral mRNAs. In experiments in which the cloning procedure involved the synthesis of a complementary DNA copy of particular *in vivo* viral mRNAs, the cloned DNAs themselves were found to contain a short stretch of nonviral nucleotides at the end corresponding to the 5′ end of the viral mRNA [6, 17, 22]. Also, by using appropriate restriction fragments derived from the cloned DNAs as primers for the reverse transcriptase-catalyzed copying of the 5′ ends of several of the *in vivo* viral mRNAs (primer extension), it was shown that these viral mRNAs contained 5′-terminal sequences that are heterogeneous and are not encoded by the vRNA [6, 17, 57, 60].

B. Priming by Cellular Capped RNAs—Mechanism

The basic mechanism of the priming reaction was elucidated in *in vitro* studies using the virion-associated transcriptase [77]. The mechanism established by these studies is shown in Fig. 1. The first step in the reaction is the endonucleolytic cleavage of the capped RNA primer at a purine residue 10 to 13 nucleotides from the cap. To demonstrate the activity of the endonuclease, several capped RNAs containing ^{32}P label only in their 5′-terminal methylated cap structure were incubated with detergent-treated virus or purified viral nucleocapsids (cores) in the

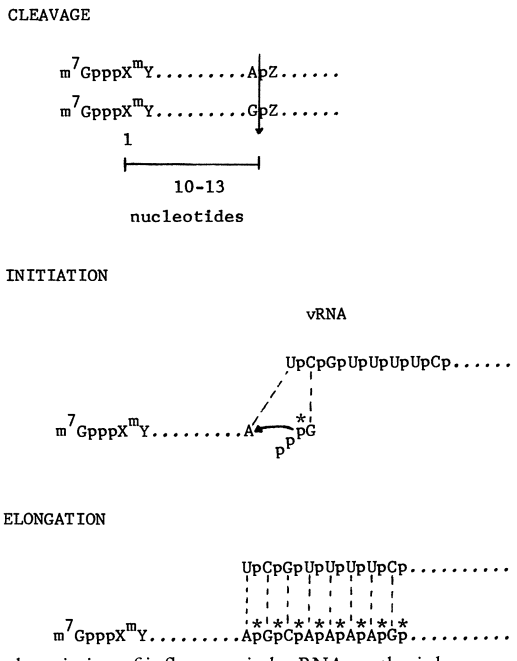


Fig. 1. Mechanism for the priming of influenza viral mRNA synthesis by capped RNAs. For description, see text. From Krug [42]

absence of ribonucleoside triphosphates [77]. For example, with brome mosaic virus (BMV) RNA segment 4 as substrate, two specific cleavages occurred (Fig. 2, lane 2): at nucleotide G12, (m^7GpppGmUAUUAAUAAUG) and at nucleotide A10. No cleavage at U11 occurred, demonstrating a strong preference for cleavage at purines. These cleavage products contained a 3′-hydroxyl group, as expected for primer molecules. The endonuclease requires the presence of a 5′-terminal methylated cap structure. Thus, for example, BMV RNA 4 with a 5′-GpppG terminus was not cleaved (Fig. 2, lane 3).

The second step in the reaction is the initiation of transcription via the incorporation of a G residue onto the 3′ end of the capped primer fragment generated by the endonuclease (Fig. 1). This intermediate in the transcription reaction can be identified by incubating an unlabeled primer RNA with the virion transcriptase complex in the presence of (α-^{32}P)GTP as the only ribonucleoside triphosphate [77]. Of the primer fragments generated by the endonuclease, those containing a 3′-terminal A residue are greatly preferred over those containing a 3′-terminal G residue in the initiation step. Thus, with BMV RNA 4, the A10 fragment is greatly preferred over the G12 fragment. When unlabeled BMV RNA 4 was incubated with detergent-treated virus in the presence of (α-^{32}P)GTP, the only primer fragment detected was the A10 cleavage product linked to one or more G residues. The utilization of the A10 fragment could also be shown using ^{32}P-cap-labeled BMV RNA 4 as primer in the presence of unlabeled GTP (Fig. 2, lane 3): the A10 fragment disappeared and was replaced by three more slowly migrating bands, which were

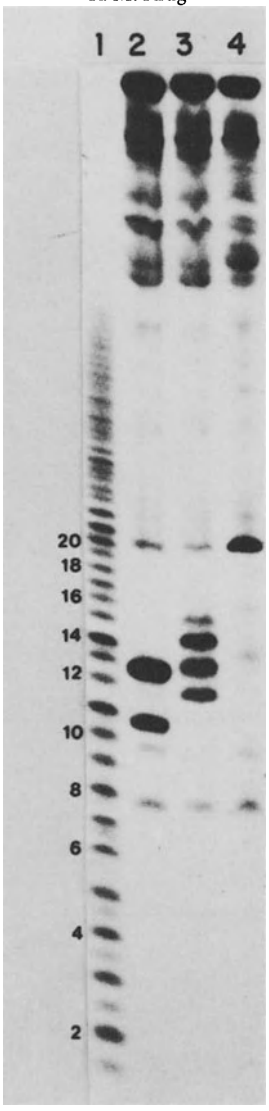


Fig. 2. Specific cleavage of BMV RNA 4 by the influenza virion endonuclease. BMV RNA 4 containing either a ^{32}P-labeled methylated cap structure m^{7}G$\overset{*}{p}$ppGm (lanes *2* and *3*) or a ^{32}P-labeled unmethylated cap structure G$\overset{*}{p}$ppG (lane *4*) was incubated with detergent treated virus in the absence of ribonucleoside triphosphates (lanes *2* and *4*) or in the presence of 25μM unlabeled GTP (lane *3*). After incubation, the phenol-extracted RNA was analyzed by electrophoresis in a 20% acrylamide gel in 7M urea. Lane *1* is the partial alkali digestion product of BMV RNA 4 containing an m^{7}G$\overset{*}{p}$ppGm cap, and the numbers on the left refer to the chain lengths of these products, counting the Gm residue as the first base. From Plotch *et al.* [77]

shown to be the A10 fragment with one, two, or three G residues added. The G12 fragment generated by the endonuclease was not used directly as a primer, but was first converted to the A10 fragment. The preference for A-terminated fragments also extends across different capped RNA species. The A-terminated fragments of BMV RNA 4 and alfalfa mosaic virus (AlMV) RNA 4 are much more effeciently utilized for initiation (G-addition) than the G-terminated fragment of β-globin mRNA.

The initial G incorporation is almost certainly directed by the penultimate C residue at the 3′ end of each vRNA template (3′ UCG ...). Thus, A-terminated primer fragments could be linked only to a G residue in transcriptase reactions containing a single ribonucleoside triphosphate [77]. No incorporation of A, U, or C occurred in single triphosphate reactions. The residue incorporated after the initial G is a C, most likely directed by the G at the third position from the 3′ end of the vRNA. Even though the 3′-terminal U of the vRNA apparently cannot direct incorporation of an A residue, the preferential usage of A-terminated primer fragments results in the presence of an A residue in the viral mRNAs at the position opposite the 3′-terminal U of the vRNA. After initiation with a G residue, chain elongation ensues (Fig. 1).

It has been established that the stimulation of the initiation of transcription by capped RNAs does not require hydrogen-bonding between the capped RNA and the 3′ end of the vRNA templates. Capped 5′-terminal fragments of β-globin mRNA and AlMV RNA 4 as short as 14 to 23 nucleotides in length and capped ribopolymers, such as capped poly (A) and capped poly (AU) (1 : 1), all of which lack sequences complementary to the common 3′-terminal sequence of the vRNA templates, are effective primers [42, 44, 47]. Consequently, the stimulation of the initiation of transcription must result from a specific interaction between the capped RNA and one or more proteins in the transcriptase complex. This interaction requires the recognition of the 5′-terminal methylated cap structure of the primer RNA.

The mechanism of initiation of viral RNA transcription in the infected cell is similar in most respects to the mechanism established for *in vitro* transcription. In cloned DNAs containing a copy of the host-donated primer sequence, the 3′-terminal U of the vRNA is not always represented by an A residue, whereas the rest of the vRNA sequence is faithfully represented [6, 17, 22, 42]. It can thus be concluded that the residue opposite the 3′-terminal U of the vRNA is not virus coded but is donated by the host cell primer and that transcription *in vivo*, like that *in vitro*, initiates with a G residue directed by the penultimate C of the vRNA [6, 60]. On the other hand, the reverse transcriptase-catalyzed copying of the 5′ ends of the entire population of a particular *in vivo* viral mRNA indicated that these viral mRNAs contain almost exclusively an A opposite the 3′-terminal U of the vRNA. This indicates that cloned DNAs containing a residue other than an A at this position represent minor species and that A-terminated cellular capped fragments are preferentially, or almost exclusively, used as primers *in vivo*, as is the case *in vitro*. The host-donated sequence is heterogeneous for most of its length, but unexpectedly was found to contain predominantly a 3′ penultimate C residue [6, 60]. Somewhat less conclusive evidence was obtained for the presence of a G

residue on the 5′ side of this C. This means that a specific subset of A-terminated capped fragments, those containing (G)CA at their 3′ ends, are preferentially used as primers *in vivo*. The preferential usage of (G)CA-terminated capped fragments was not observed in earlier *in vitro* studies, as none of the capped RNAs used in these *in vitro* studies contained a (G)CA sequence 10 to 13 nucleotides from the cap. At present, the basis for the preferential usage of (G)CA-terminated host capped fragments *in vivo* is not known. Nor is it known whether this population of host cell primers is comprised of a large or relative small number of different species.

C. Role of the Three Viral P Proteins in the Steps of Primed Transcription

The entire process of capped RNA-primed transcription is catalyzed by purified viral cores [77, 99], which contain the vRNA segments and four known virus-specific proteins, the nucleocapsid protein (NP) and the three P proteins [36, 77, 99]. The NP protein, which constitutes the majority (about 92%) of the protein in viral cores, probably has primarily a structural role, as it is situated along the vRNA chains at approximately 20-nucleotide intervals [20]. The P proteins, the presumed subunits of the transcriptase, distribute during two-dimensional gel electrophoresis into two basic species, PB1 (the larger of the two basic P proteins) and PB2 (the smaller basic P protein), and one acidic species, PA [34, 99] (Fig. 3 A). Several approaches have been employed to establish the role of each of the P proteins in individual steps of viral mRNA synthesis.

Crosslinking experiments have been used to identify the P protein which is closely associated with cap structures and hence is likely to be the viral cap-recognizing protein. In one approach, the capped RNA primer itself, labeled with ^{32}P only in its cap structure, was used as a radioactive probe for the cap-recognizing protein [99]. Endonuclease reactions carried out by viral cores using this cap-labeled RNA primer (A1MV RNA 4) in the absence of ribonucleoside triphosphates, were irradiated with ultraviolet (UV) light. After nuclease digestion, the labeled cap was found to be crosslinked to a protein which has a mobility similar to that of the PB2 protein in two-dimensional gel electrophoresis (Fig. 3 B). The crosslinked protein and unreacted PB2 protein differed slightly in mobility, presumably because the former contained a few covalently-linked nucleotides. The specificity of the crosslinking of PB2 to the labeled cap of the primer RNA was demonstrated by competition experiments in which only unlabeled RNAs that contained a cap 1 structure blocked the crosslinking. In another approach, a derivative of m^7GTP photoreactive at 320 nm was used as an affinity label [8]. One of the viral P proteins became crosslinked to this probe and this crosslinking was competitively and specifically inhibited by cap analogues. Initially, the identity of the crosslinked P protein was not clearly established, because the viral proteins were analyzed only by one-dimensional gel electrophoresis [8]. Subsequent experiments employing two-dimensional gels showed that the crosslinked P protein is in fact PB2 [9], thereby confirming the results obtained with the cap-labeled primer RNA and UV crosslinking [99].

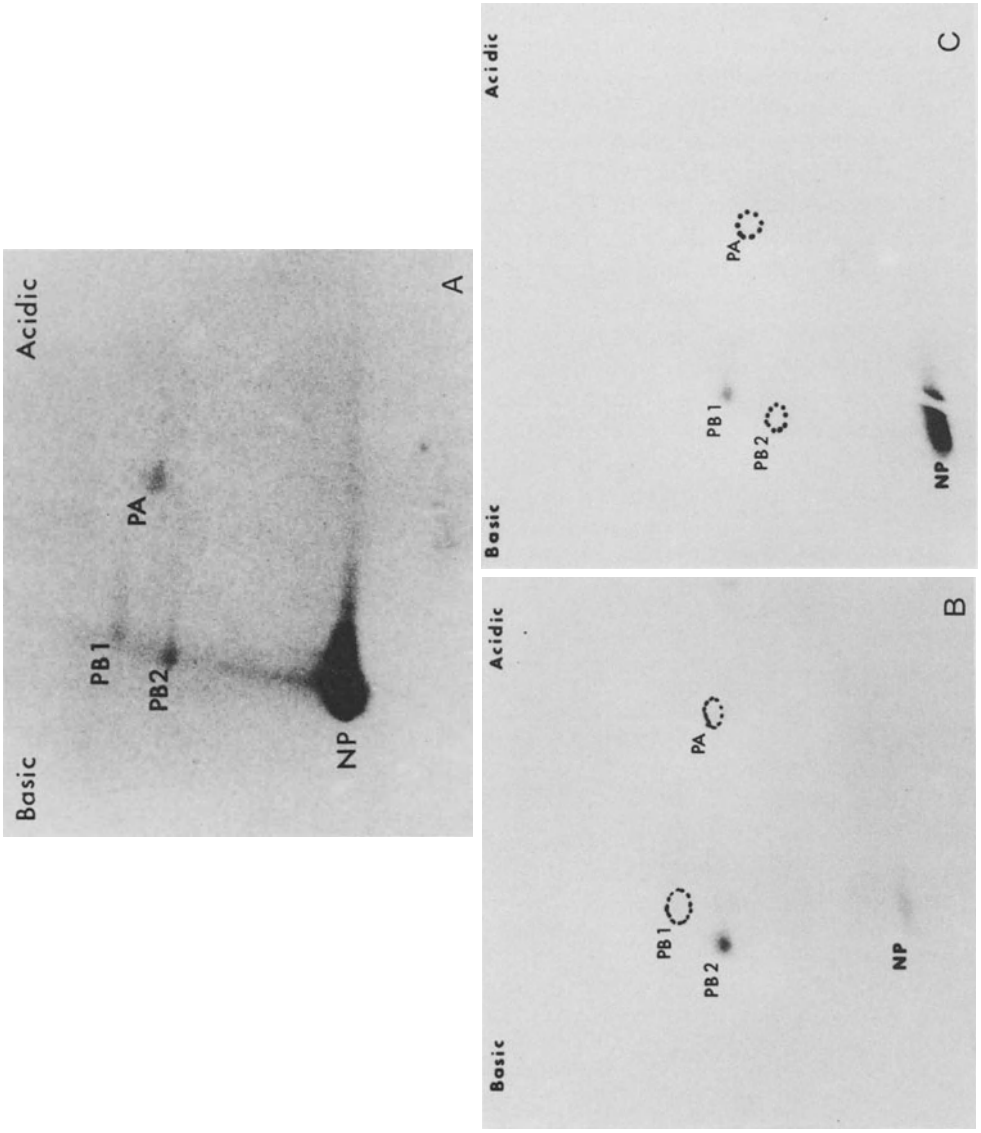


Fig. 3. Two-dimensional gel electrophoresis of ^{35}S methionine-labeled proteins in viral cores *(A)*, of the crosslinked ^{32}P-labeled cap 1-recognizing core protein *(B)*, and of the crosslinked ^{32}P-labeled core protein which most likely catalyzes the initiation of transcription. For B, purified viral cores were incubated for 30 minutes at 31C with ^{32}P-cap-labeled A1MV RNA 4 in the absence of ribonucleoside triphosphates. For C, purified viral cores were incubated for 30 minutes at 31C with capped globin mRNA in the presence of (α-^{32}P)GTP as the only labeled ribonucleoside triphosphate. Reactions B and C were exposed to UV light. These two reaction mixtures and also the ^{35}S-methionine-labeled viral cores *(A)* were digested with pancreatic and T1 RNases, and were analyzed on two-dimensional gels as described in Ulmanen *et al.*, [98]. The expected positions of the PB1 and PA proteins are marked in B, and the expected positions of the PA and PB2 proteins are marked in C. From Ulmanen *et al.* [99]

These crosslinking experiments demonstrate a topographical relationship between PB2 and the cap 1 structure, but do not firmly establish that the PB2 protein actually functions in cap recognition during viral RNA transcription. Definitive evidence for a function of PB2 in cap recognition has been provided by a recent study of the temperature sensitivity of the *in vitro* transcription reaction catalyzed by virus mutants with a temperature-sensitive *(ts)* defect in the genome RNA segment coding for the PB2 protein [100]. The two mutants used, WSN mutants tsl and ts6, behave as *ts* transcription mutants *in vivo* [49]. To obtain interpretable data about the temperature sensitivity of *in vitro* transcription catalyzed by these mutants, it was first necessary to overcome a major difficulty that plagued previous studies: the apparent temperature sensitivity of the transcriptase activity of wild type *(wt)* virus *in vitro*. At the temperatures (>39 C) which are nonpermissive for *ts* virus mutants *in vivo*, the transcriptase of detergent-treated *wt* virions exhibits *in vitro* only about 10% of the activity seen at the optimal *in vitro* temperature (30–33 C) [72, 74]. Consequently, investigators had attempted to determine whether the transcriptase activity associated with a given *ts* virus mutant exhibited even less than 10% of the optimal activity at these higher temperatures (>39 C). For this reason, the data and their interpretation had been less than conclusive [71, 72,

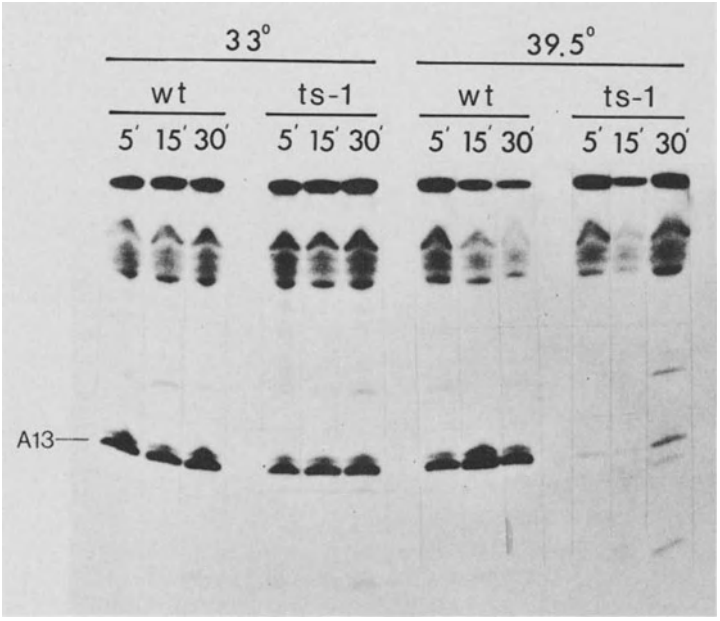


Fig. 4. The effect of temperature on the endonuclease reaction catalyzed by purified viral cores from wt and tsl virus. Viral cores from wt or tsl virus were incubated with ^{32}P cap-labeled AlMV RNA 4 for the indicated times at 33 or 39.5 C under endonuclease assay conditions (in the absence of ribonucleoside triphosphates). The labeled cleavage products were analyzed on 20% acrylamide/7 M urea gels. The position of the A13 fragment is indicated. The labeled RNAs migrating between the intact AlMV RNA 4 (which remained at the origin) and the A13 cleavage product are partial degradation products of AlMV RNA 4 present in the unreacted ^{32}P AlMV RNA 4 preparation. These RNAs are also substrates for the endonuclease, as they along with the full-length AlMV RNA 4 are cleaved when an excess of wt viral cores are used in the reaction. From Ulmanen *et al.* (100)

74]. This difficulty was overcome by using purified viral cores rather than detergent-treated virions to catalyze transcription [100]. Detergent-treated virions contain undefined destabilizing and/or inhibitory activities which are more active at higher temperatures, and these activities are largely, if not entirely, removed during the purification of viral cores. Using purified viral cores of *wt* WSN virus, the rate of endonucleolytic cleavage of capped RNA primers and of overall transcription was shown to be similar at 39.5 C and 33 C, the *in vivo* nonpermissive and permissive temperatures, respectively, of the WSN *ts* mutants. Thus, in endonuclease assays using cap-labeled A1MV RNA 4, *wt* viral cores generated approximately the same amount of the specific A13 cleavage product at each time point at 33 C and 39.5 C (Fig. 4). In contrast, ts1 viral cores, which have activity similar to that of *wt* at 33 C, catalyzed very little specific cleavage at 39.5 C [100]. The ts6 viral cores also exhibited a *ts* cap-dependent endonuclease. The steps in transcription after the endonucleolytic cleavage of the capped RNA primer were shown to be essentially temperature-insensitive, indicating that the mutations in the PB2 protein found in ts1 and ts6 virions affect only the endonuclease step. The *ts* defect is most likely in the recognition of the 5′-terminal cap 1 structure that occurs as a required first step in the endonuclease reaction: the cap-dependent binding of a specific capped primer fragment (the A13 fragment of A1MV RNA 4) to ts1 viral cores was temperature-sensitive under conditions in which binding to *wt* viral cores was not affected by increasing the temperature from 33 to 39.5 C [100]. These results thus establish that the viral PB2 protein does in fact function in cap recognition during the endonuclease reaction.

UV crosslinking experiments have also been used to identify which viral P protein(s) catalyze the initiation of transcription and which protein probably catalyze chain elongation (Fig. 1). To identify the P protein that catalyzes the initiation of transcription via the incorporation of a G residue onto the end of primer fragments, viral cores were incubated with an unlabeled primer RNA (*e.g.*, globin mRNA) and (α-^{32}P)GTP as the only ribonucleoside triphosphate and the reaction mixtures were UV-irradiated [99]. After nuclease treatment, the labeled G residue was found to be crosslinked to a protein which has a mobility similar to that of the PB1 protein in two-dimensional gel electrophoresis (Fig. 3 C). As with the PB2 protein, a small difference was detected between the mobility of the crosslinked PB1 protein and that of the unreacted PB1 protein, consistent with the presence of a few covalently linked nucleotides after crosslinking and nuclease digestion. The transcriptase reaction conditions required for bringing the PB1 protein into close association with a labeled G residue so that crosslinking can occur are the same conditions required for covalent bond formation between the G residue and the endonuclease-generated primer fragment. This strongly suggests that the close association of the labeled G residue with the PB1 protein occurs coincident with incorporation of G onto the primer fragment and that the PB1 protein catalyzes the initiation step of transcription [99]. Pulse-chase experiments carried out in the author's laboratory indicate that soon after chain elongation the PB1 protein is no longer associated with the initiating G residue (Braam, Ulmanen, and Krug, unpublished experiments). Elongation may also be catalyzed at least in part by the PB1 protein. In initial experiments in which elongation was interrupted to generate

chains of varied lengths, it was found that the last residue added to the growing viral mRNA chains could be crosslinked to, and hence is apparently closely associated with, the viral PB1 protein (Braam, Ulmanen, and Krug, unpublished experiments).

These results begin to suggest a coherent model for the actions of the individual viral P proteins during viral mRNA synthesis. However, several important questions remain. For example, it is not yet known which viral P protein catalyzes the cleavage of the capped RNA at a purine residue 10 to 13 nucleotides from the cap. In other words, does the PB2 protein not only recognize the 5′-terminal cap structure of the primer RNA but also cleave this RNA 10 to 13 nucleotides from the cap, or is the actual cleavage catalyzed by a second P protein, such as PB1? Also, if it is verified that PB1 elongates the viral mRNA chains as it moves down the template vRNA, does the PB1 protein recycle to elongate viral mRNA chains on the same or a different vRNA nucleocapsid template? Further understanding of the functions and movements of the three P proteins during viral mRNA synthesis may be expected to provide insights into the way viral mRNA synthesis is regulated and into the mechanism of the switch from viral mRNA to full-length transcript synthesis (see later).

D. Termination of Transcription and Poly (A) Addition

Both the *in vivo* and *in vitro* viral mRNAs contain 3′-terminal poly (A) sequences. The poly (A) sequences found in the *in vivo* viral mRNAs are shorter and more homogeneous in length than those found in the *in vitro* viral mRNAs [78]. Experiments in which hybrids between vRNA segments and poly (A)-containing viral mRNAs were digested with single-strand specific nucleases indicated that the viral mRNAs lack the sequence complementary to the 5′ terminal 20 to 30 nucleotides of the vRNAs [29, 79]. In this region of each of the eight vRNA segments (about 17 to 22 nucleotides from the 5′ end), is a tract of 5 to 7 U residues [83, 93]. This short U tract has been established as the site of poly (A) addition [84]. To demonstrate this, an individual vRNA segment was digested with RNase T1, and the resulting oligonucleotides after 5′ labeling were annealed to viral mRNA. After T1 and pancreatic RNase digestion, the labeled vRNA T1 oligonucleotide that was base-paired to the poly (A)-containing mRNA obligonucleotide was isolated by (oligo dT)-cellulose chromatography and was sequenced. The vRNA T1 obligonucleotide isolated by this procedure was found to contain the U tract located 17 to 22 nucleotides from the 5′ end of the vRNA.

Poly (A) addition at this site could occur by two mechanisms: (i) endonucleolytic cleavage of a larger transcript (*i.e.*, one containing a copy of the 5′ terminal 17 to 22 nucleotides of the vRNA), followed by polyadenylation; or (ii) "stuttering" or repetitive copying of the short 5 to 7 U tract of vRNA and consequent termination of copying of the ensuing 5′ terminal vRNA sequences. Endonucleolytic cleavage of a larger transcript has been established as the mechanism of poly (A) addition for some RNA polymerase II transcription units [73], while recent evidence suggests that for other polymerase II transcription units termination of transcription may

operate [66]. For influenza viral RNA transcription, it has been presumed that repetitive copying of the U tract and termination of transcription, rather than endonucleolytic cleavage of a larger transcript, occurs. No evidence for transcription of the 5′-terminal sequences of vRNA during *in vitro* viral mRNA synthesis (catalyzed by the virion-associated transcriptase) has been reported. It is known what features of the template could signal poly (A) addition and consequent transcription termination. Two other negative-strand viruses, vesicular stomatitis virus (VSV) and Sendai virus, have a consensus tetranucleotide—AUAC and AUUC, respectively—immediately preceding the U tract at which poly (A) addition occurs, and this tetranucleotide has been postulated to be the signal for poly (A) addition [26, 90]. In contrast, the influenza vRNA segments do not contain a common sequence preceding the U tract [83, 93]. As an alternative, it has been proposed that the partial sequence complementarity between the 5′ and 3′ ends of each individual vRNA segment may result in the formation of a duplex "panhandle" structure extending up to the U tract, thereby blocking transcription of the 5′ terminal vRNA nucleotides and fostering repetitive copying of the U tract [84].

E. Regulation of Viral mRNA Synthesis in the Infected Cell

The synthesis of influenza viral mRNAs is controlled during infection with respect to the relative amount of each mRNA synthesized and to the time at which each mRNA is synthesized in greatest amount. This regulation has been most intensively studied in chick embryo fibroblasts infected with fowl plague virus [29]. Immediately after infection, similar amounts of all eight viral mRNAs were detected. This was followed by a period, often termed the early phase, during which the synthesis of two viral mRNAs predominated, those coding for the NP protein and for the nonstructural protein, NS1. Subsequently, the rate of synthesis of the NS1 mRNA fell relative to that of the NP mRNA, and the rate of synthesis of the M (membrane protein), HA (hemagglutinin) and NA (neuraminidase) mRNAs greatly increased. This phase has been termed the late phase of viral mRNA synthesis. At all times except immediately after infection, the rate of synthesis of the three P protein mRNAs remained low relative to that of the other viral mRNAs. This basic overall pattern, possibly with some variations, likely occurs in other productive virus infections. The feature that has been taken as most representative of the regulation of viral mRNA synthesis is the shift in the relative rates of synthesis of the NS1 and M mRNAs, a shift from a predominance of NS1 mRNA synthesis (early phase) to predominance of M mRNA synthesis (late phase). The relative rates of synthesis of the various viral mRNAs at different times of infection correlates closely with the relative rates of synthesis of the proteins encoded by these mRNAs [29, 37, 51, 69], indicating that viral gene expression is regulated largely at the level of transcription.

The mechanism by which viral mRNA synthesis is regulated during infection is not known. Primary transcription catalyzed by the inoculum virus is not regulated: similar relative amounts of the different viral mRNAs are synthesized immediately after infection and also in cells infected in the presence of cycloheximide [4, 29]. In

addition, the virion-associated transcriptase *in vitro* does not exhibit the regulation seen during infection. Rather, in the *in vitro* reaction the smaller molecular weight mRNAs, the NS1 and M mRNAs, are synthesized in greatest amount and decreasing amounts of the larger size mRNAs are made [79]. This probably only reflects the fact that the *in vitro* reaction is not optimized for chain elongation, so that smaller size chains are made in largest amounts. Recent experiments in the author's laboratory suggest that the different viral mRNA chains are initiated at similar rates (Braam and Krug, unpublished experiments). Consequently, the regulation of viral mRNA synthesis is most likely mediated by virus-specific products synthesized after infection. It has been suggested that the NS protein(s) play a role in this regulation. This was suggested because the shift from the early to the late pattern of viral protein synthesis (and hence presumably of viral mRNA synthesis) was inhibited in cells infected at the nonpermissive temperature by *ts* virus mutants containing a defect in the genome segment coding for the NS proteins [39, 104]. It is not clear exactly how the NS protein(s) might regulate the switch from the early to late transcription pattern. Also, since the early regulation of viral mRNA synthesis occurs at the nonpermissive temperature, this early regulation would have to be attributed to a virus-coded component other than the NS protein(s). Moreover, with two of the NS *ts* virus mutants, vRNA synthesis was also inhibited at the nonpermissive temperature [104], so that the block in the shift from the early to the late phase of viral mRNA synthesis could actually be a consequence of the block in vRNA synthesis. In this regard, recent evidence suggests that the regulation of viral mRNA synthesis may occur at least in part at the level of vRNA synthesis [94]. It was found that, particularly at early times of infection, there was a good correlation between the relative rates of synthesis of the different vRNA segments and their encoded mRNAs. On this basis, it was postulated that viral mRNA itself is not regulated, but rather that the relative amount of synthesis of a particular viral mRNA at any time is proportional to the amount of vRNA synthesized at that time. According to this hypothesis, a given vRNA template would only be transcribed into mRNA for a short time; or in fact, a given vRNA may only be transcribed once into mRNA. A problem with this theory is that viral mRNA synthesis is apparently not dependent on ongoing vRNA synthesis. When vRNA synthesis is blocked by a protein synthesis inhibitor (cycloheximide), viral mRNA synthesis continues—in fact, often at an accelerated rate [25, 51, 67]. As an answer to this objection, it was proposed that inhibiting protein synthesis somehow "uncouples" viral mRNA synthesis from ongoing vRNA synthesis [93].

F. Cellular Site of Viral mRNA Synthesis

For many years, investigators attempted to determine whether influenza viral mRNA synthesis occurs in the nucleus or cytoplasm of infected cells. The early results were inconclusive and conflicting. Some evidence was consistent with a nuclear site; for example, the presence of internal m^6A residues in the viral mRNA and the likely generation of a few of the viral mRNAs by splicing [43, 46, 50]. However, studies of viral RNA-dependent RNA polymerase activity (synthesizing

RNA complementary to vRNA) in infected cells yielded ambiguous results. A few studies detected such an activity in isolated nuclei [27, 65], but in most studies activity was found in the cytoplasmic fraction of infected cells [15, 19, 63, 65, 88, 92]. Whether any of these activities is actually responsible for viral RNA transcription *in vivo* is unclear. In fact, it is not likely that the cytoplasmic activity is functional *in vivo,* because most of this activity was detected late in infection [19, 63], after most viral RNA transcription in the cell was completed [4, 29, 67, 89, 98]. The cytoplasmic activity may be associated with progeny nucleocapsids that are about to be packaged into virions. Other studies also did not provide a clear-cut answer. Experiments measuring the steady-state levels of viral RNA transcripts indicated that these transcripts accumulate predominantly in the cytoplasm [4, 67, 98]. However, these experiments do not have the capability of definitively identifying the site of synthesis of the transcripts that accumulated in the cytoplasm. Nonetheless, one result from these studies was interpreted as suggesting that primary transcription occurs in the nucleus. Finally, one group of investigators reported pulse-labeling experiments that pointed to the cytoplasm as the site of viral RNA transcription, both primary and amplified. Using ^{3}H-uridine pulses of 5 to 30 minutes, Hay *et al.* [29] found the majority of pulse-labeled viral RNA transcripts in the cytoplasm of infected chicken embryo fibroblasts (CEFs).

The site of viral mRNA synthesis was resolved by a recent study [33]. To measure pulse-labeled viral RNA transcripts, pulse-labeled infected cell RNA was hybridized to vRNA covalently bound to filter paper discs. Both aqueous and nonaqueous cell fractionation methods were used with two different cell types, CEFs and BHK-21 cells, infected with the virus, to establish conclusively that the location of pulse-labeled viral RNA transcripts after fractionation of cells accurately reflected the location of these molecules in intact cells. The distribution of two cellular RNAs, 45S ribosomal precursor RNA (a nuclear RNA) and mature 18S ribosomal RNA (a cytoplasmic RNA), were used to monitor the purity of the nuclear and cytoplasmic fractions. This study demonstrated that viral RNA transcription occurs in the nucleus. The principal difficulty encountered was that the influenza viral RNA transcripts (which are 12 to 18S in size) leaked out of the nuclei of CEFs during aqueous fractionation, whereas the larger RNA (45S ribosomal precursor RNA), used as a marker for nuclear leakage, remained in the nuclei. With CEFs, a substantially larger percentage of the pulse (6 minute)-labeled viral RNA transcripts was found in the nucleus after nonaqueous fractionation (55—59%) than after aqueous fractionation (15—23%). This finding probably accounts for the results of Hay *et al.* [29], who employed aqueous fractionation of CEFs and found pulse-labeled viral RNA transcripts predominantly in the cytoplasmic fraction. In contrast with the case in CEFs, pulse-labeled viral RNA transcripts were retained in the nuclei of BHK-21 cells during both aqueous and nonaqueous fractionation [33]. With a 2 minute pulse of ^{3}H-uridine at 1.5 hours after infection of BHK-21 cells, 78—90% of the labeled viral RNA transcripts were found in the nuclear fraction. After this study, another group obtained evidence indicating that viral RNA transcription occurs in the nucleus [38]. These investigators showed that viral RNA transcripts pulse-labeled with ^{3}H-uridine for 2.5 minutes are associated with a subnuclear structure, termed the nuclear cage, the

site at which newly synthesized cellular RNA transcripts are also apparently found.

Viral RNA transcription most likely occurs in the nucleus throughout infection. Most viral RNA transcription occurs during the first 2.5 to 3 hours after infection [4, 29, 67, 89, 98], and pulse-labeled viral RNA transcripts were shown to be predominantly in the nucleus at the two time points analyzed (1.5 and 2.5 hours) during this period of maximum viral RNA transcription [33]. Pulse-labeled viral RNA transcripts were also found predominantly in the nucleus under conditions (inhibition of protein synthesis) in which only primary transcription occurs. It was shown that the pulse-labeled viral RNA transcripts in the nucleus are comprised

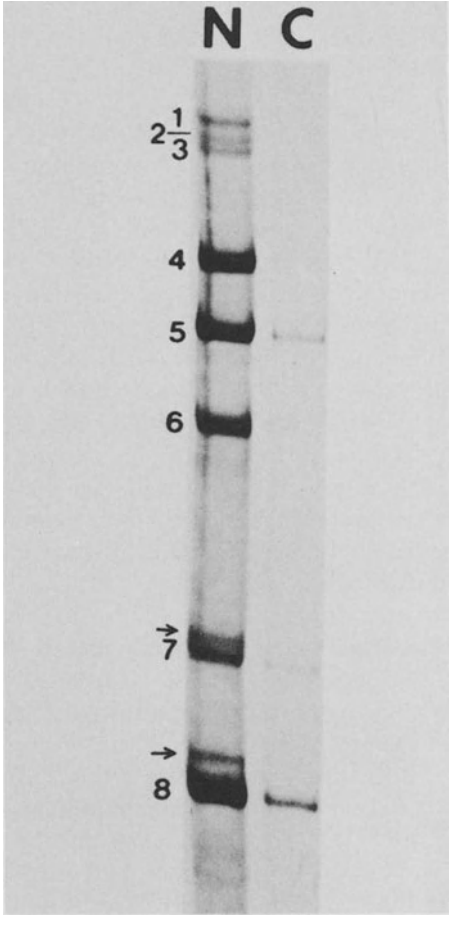


Fig. 5. The newly synthesized (pulse-labeled) segments of influenza viral positive stranded RNA (mRNA and full-length transcripts) in the nucleus *(N)* and cytoplasm *(C)* of infected BHK-21 cells. At 1.5 hours after infection, the BHK-21 cells were labeled with ^{3}H-uridine for 6 minutes. The cells were fractionated by a nonaqueous procedure, and the nuclear and cytoplasmic RNAs were annealed to excess unlabeled vRNA. After treatment with RNase T 2, the resulting hybrids were resolved by gel electrophoresis. The eight vRNA: ^{3}H-labeled viral mRNA hybrids are indicated on the left. Arrows: hybrids of vRNA: ^{3}H-labeled full-length transcripts. From Herz *et al.* (33)

primarily of the eight viral mRNAs [33] (Fig. 5). Since the viral mRNAs function in the cytoplasm, it was important to demonstrate that the newly synthesized viral mRNAs detected in the nucleus are in fact transported to the cytoplasm. Pulse-chase experiments with ^{3}H-uridine using glucosamine to deplete the uridine pool provided strong evidence for transport [33]. In the BHK-21 cells used for these experiments, the chase conditions were not totally effective in stopping ^{3}H-uridine incorporation, but the changing nuclear-cytoplasmic distribution of labeled viral RNA transcripts during the chase period indicated that transport of a large proportion of these transcripts most likely occurs. Thus, the proportion of labeled viral RNA transcripts found in the nucleus decreased from 91% after a 4 minute ^{3}H-uridine pulse (at 1.5 hours after infection) to 31% after a 60 minute chase.

The viral mRNAs synthesized in the nucleus must therefore be primed by capped host cell nuclear RNAs. This was verified directly using ^{3}H-methyl methionine to label the newly synthesized cap 1 structures of influenza viral mRNA: the majority of the pulse-labeled caps were found in the nucleus [33]. Therefore, the cellular primers actually used for viral mRNA synthesis are derived from hnRNAs. It is not known whether the 5′ ends of cellular hnRNAs are cleaved by the viral endonuclease to generate primer fragments when the hnRNA chains are nascent, completed, or partially and/or completely processed. Nascent hnRNA chains might be used, because the 5′-terminal cap structure required by the viral endonuclease is added to nascent hnRNA chains soon after chain initiation [86]. The use of hnRNAs as primers explains why only capped RNAs synthesized after infection are used as primers, and consequently why α-amanitin added at the beginning of infection totally inhibits viral RNA transcription [67]. Direct evidence for the use of only newly synthesized capped cellular RNAs as primers has been obtained: the 5′ methylated caps on the viral mRNAs were labeled by exposing cells to ^{3}H-methyl methionine only after, and not before, infection [33]. The observed inhibition of viral RNA transcription when α-amanitin is added after infection [67] can be attributed to the necessity for RNA polymerase II to continue to synthesize hnRNA primers. Continued synthesis may be necessary because: (i) hnRNAs are being rapidly depleted not only by normal turnover and transport (after processing) to the cytoplasm, but also by their continuous utilization as primers for viral mRNA synthesis; or (ii) only nascent hnRNAs are used to generate primers. It is not known what fraction of the newly synthesized hnRNAs in infected cells is used for the generation of primers for viral mRNA synthesis. If a significant fraction of the hnRNAs is used, one would expect clearcut effects on cellular hnRNA metabolism as a result of virus infection.

The demonstration of a nuclear site for viral mRNA synthesis explains earlier observations that the viral NP and P proteins are transported to the nucleus [13, 30, 45, 61, 97], where they are found in the form of viral nucleocapsids containing vRNA [40, 41]. These nucleocapsids remain in the nucleus and are not precursors of the viral nucleocapsids in the cytoplasm that are incorporated into virus particles [41]. The time course of appearance of the nuclear (nucleoplasmic) viral nucleocapsids is similar to the time course of viral mRNA synthesis [41]. It is thus likely that these nucleoplasmic viral nucleocapsids are the transcriptase complexes

responsible for viral mRNA synthesis in the infected cell. Indeed, recent experiments in the author's laboratory have detected in nuclear preparations from infected cells at early times (as early as two hours) an activity that catalyzes the synthesis of poly (A)-containing, virus-specific transcripts with the same electrophoretic mobilities as the eight major viral mRNAs (Beaton and Krug, unpublished experiments). In contrast with the virion-associated transcriptase, the nuclear transcriptase exhibits substantial activity in the absence of an added capped RNA primer, indicating that viral mRNA chains already initiated *in vivo* are being elongated *in vitro* and/or that cellular capped RNAs which accumulated in the nuclei *in vivo* are priming the *in vitro* reaction. The nuclear transcriptase is also capable of initiating viral mRNA chains *in vitro,* because the addition of a capped RNA stimulates viral mRNA synthesis 2 to 5-fold (rather than the almost 100-fold seen with the virion-associated transcriptase) and the 5′ end of the added capped RNA is incorporated into the resulting viral mRNAs.

In addition to the NP and P proteins, another virus-coded protein—the NS1 protein—accumulates in the nucleus [30, 45, 48, 61, 97]. This protein is found in both the nucleolar and nucleoplasmic fractions [45, 48]. Recent experiments indicate that the distribution of the NS1 protein between the nucleolus and nucleoplasm varies with different virus strains and different cell types (Young and Palese, unpublished experiments). The function of the NS1 protein in the nucleus is not known.

G. Role of Other Host Nuclear Functions in Viral mRNA Synthesis— Splicing and Methylation of Internal A Residues

Influenza virus apparently utilizes other host cell nuclear functions during the synthesis of its viral mRNAs. The viral mRNAs, like other mRNAs synthesized in the nucleus, contain internal m^6A residues [44, 46], and it is likely that this methylation is catalyzed by host nuclear enzymes. In addition, strong evidence has been obtained indicating that two of the viral mRNAs, those coding for the M1 and NS1 proteins, are themselves spliced to form smaller mRNAs [50]. As the splice junctions in these mRNAs are similar to those found in RNA polymerase II-catalyzed transcripts, it is likely that cellular enzymes are used for splicing these viral mRNAs. If verified, these would be the only known examples of nuclear splicing of RNA transcripts that are not DNA-directed and that are not synthesized by RNA polymerase II.

The discovery of splicing stemmed from the identification in influenza virus A-infected cells of a second nonstructural protein, NS2, molecular weight (11,000), which was shown to have a tryptic peptide pattern different from those of the other eight viral gene products. This meant that one of the eight vRNA (genome) segments had to code for two proteins [56]. Segment 8, the smallest vRNA segment, which codes for the NS1 protein, was shown to be the segment which also codes for the NS2 protein. This was established by studies with recombinant viruses in which the NS1 and NS2 proteins were shown to reassort together and by hybrid-

arrest translation experiments [35, 53]. Two different mRNAs code for the NS1 and NS2 proteins, and the translation of both these mRNAs *in vitro* was specifically blocked by hybridizing genome segment 8 to the total *in vivo* viral mRNA used for translation. By S1 nuclease mapping it was shown that the body of the NS1 mRNA maps from 0.05—0.95 units of the cloned NS DNA and the body of the NS2 mRNA maps from 0.59—0.95 units [55]. This suggested that the two mRNAs are 3′ co-terminal. To establish precisely which nucleotide sequences of the NS genome segment are represented in the NS2 mRNA, the 5′ end of this mRNA was sequenced by the primer extension method, using a DNA primer from the body region of the NS2 mRNA [57]. The data showed that the NS1 and NS2 mRNAs have common 5′ ends, including the host-derived primer sequence and ~56 virus-coded nucleotides. This leader sequence contains the initiation codon for protein synthesis and coding information for nine amino acids which would be common to the NS1 and NS2 proteins. Following this leader sequence in the NS2 mRNA is an interrupted region of 473 nucleotides, and the leader sequence is covalently linked to the body of the NS2 mRNA at nucleotide 526—529 (numbering system relative to genome RNA). The exact nucleotide at which joining occurs could not be determined because of the repetition of a sequence—CAGG—at positions 54—57 and 526—529 (see below). This joining continues the protein coding region of the NS2 mRNA, but the body of this mRNA (340 nucleotides) would be translated in the +1 reading frame relative to the NS1 mRNA. A schematic representation of the arrangement of the NS1 and NS2 mRNAs of influenza A virus is shown in Chapter 2, Fig. 9. Subsequently, it was shown that with influenza B viruses genome segment 8 also encodes two mRNAs [14]. Again, the NS2 mRNA contains an interrupted region, in this case 655 nucleotides in length.

Genome segment 7 of influenza A viruses also encodes more than one mRNA, in this case 3 mRNAs, and two of these mRNAs contain interrupted regions. It had been established that this genome segment codes for the M protein, a viral structural protein [75]. When the cloned DNAs of genome segment 7 from several influenza A virus strains were sequenced, it was noted that this segment contains a second open reading frame for translation [1, 58, 103]. This led to the discovery of a nonstructural protein of molecular weight ~15,000, called the M2 protein, which is encoded in a separate mRNA [54]. The *in vitro* translation of this mRNA as well as of the M (renamed M1) mRNA was blocked by hybridizing genome segment 7 to the total *in vivo* viral mRNA used for translation, indicating that the M1 and M2 mRNAs are both transcribed from genome segment 7 [60]. In addition, a third potential mRNA encoded in segment 7 was discovered [60]. This mRNA has been denoted as $mRNA_3$ rather than as M3 mRNA because a M3 polypeptide has not yet been identified. The $mRNA_3$ was discovered when the 5′ sequences of the small-size *in vivo* mRNA fraction were determined by primer extension using a DNA primer from the body of the M2 mRNA: two extension products were observed, one corresponding to M2 mRNA and one to $mRNA_3$. Sequence analysis of the M2 mRNA extension product indicated that this mRNA has the same 5′ end as the M1 mRNA, including the host-derived primer sequence and ~51 virus-coded nucleotides. There is then an interrupted region of 689 nucleotides, and the leader is then linked to the body of the M2 mRNA at nucleotide 740 or 741. The body region of

the M2 mRNA would be translated in the +1 reading frame relative to that of the M1 mRNA. The $mRNA_3$ has a shorter 5′ leader sequence, comprising only ~11 virus-specific nucleotides (plus the host-coded primer sequences), and this leader sequence is joined to the same body sequence as the M2 mRNA. The coding potential of the $mRNA_3$ is for only nine amino acids, identical to those at the carboxy terminus of the M1 protein. A schematic representation of the arrangement of the M1 and M2 mRNAs and of the $mRNA_3$ is shown in Chapter 2, Fig. 7.

There are several possible mechanisms for the generation of these viral mRNAs containing interrupted regions: (i) splicing like that occurring with DNA-directed RNA transcripts; (ii) transcriptional jumping during mRNA synthesis; and (iii) transcription of DI (defective interfering) vRNAs containing internal deletions. The data strongly favor splicing. All DI vRNA species so far characterized have been derived from the three largest genome segments coding for the three P proteins [21], and no DI vRNAs derived from segments 7 and 8 have been detected. Strong evidence for splicing comes from the sequences at the junctions between the conserved (exon) and deleted (intron) regions in the mRNAs [14, 57, 60]. Though the exact nucleotides at which the joining of the leader and body sequences of the NS2 and M2 mRNAs and of $mRNA_3$ could not be determined because of the repetition of short sequences (see above), it is clear that the sequences at the junctions between the conserved (exon) and deleted (intron) regions are similar to the consensus sequence found at both sides of the intervening regions in spliced eukaryotic mRNAs [70] (Table 1). All introns begin with GT and end with AG, and all the junctions in the influenza viral mRNAs conform to this role. The other nucleotides in the consensus sequence are not universally present in all eukaryotic splice junctions. Many, but not all, of these consensus nucleotides are found in the influenza viral mRNA junctions. For example, the acceptor sequence in the NS (NS2) mRNA of influenza A virus is a perfect match with the extended consensus acceptor sequence. On the other hand, the acceptor sequences in the NS (NS2) mRNA of influenza B virus and in the M2 mRNA and $mRNA_3$ of influenza A virus lack the pyrimidine tract adjacent to the 3′-terminal AG of the intron. Two nucleotides in the influenza viral mRNA junctions, the C at the 3′ end of the leader (5′ exon) in the M2 mRNA and the T at the 5′ end of the 3′ exon in the NS (NS2) mRNA of influenza B virus, occur only very rarely at these positions in the junctions of eukaryotic mRNAs. On the whole, however, the similarity between the influenza viral mRNA junctions and the consensus splice junctions is good.

To establish that the influenza viral mRNA junctions can be used by cellular splicing enzymes, Lamb and Lai [59] cloned m DNA (the double-stranded copy of genome segment 7) into a simian virus 40 (SV40) vector using the SV40 late promoter. A mRNA containing uninterrupted influenza virus M sequences and also both 5′ and 3′ SV40 sequences was synthesized and translated to yield authentic M1 protein. In addition, a mRNA containing an interrupted influenza virus M sequence was synthesized. The nucleotide sequences found at the junctions of this interrupted mRNA are identical to those of the $mRNA_3$ synthesized in influenza virus-infected cells. The NS DNA was also recently cloned into the same SV40 vector: uninterrupted NS1 and interrupted NS2 mRNAs were made and translated into the NS1 and NS2 proteins (Lamb and Lai, unpublished experi-

Table 1. *Comparison of the consensus splice junctions with the sequences at the junctions of the conserved (Exon) and deleted (Intron) regions of the M and NS influenza viral mRNAs*

	Donor (5′)[c]	Acceptor (3′)[c]
Consensus[a]	$\begin{matrix}C\\A\end{matrix}$AG | GT$\begin{matrix}A\\G\end{matrix}$GT	$\begin{matrix}T\\C\end{matrix}$ $_n$N$\begin{matrix}C\\T\end{matrix}$AG | G
Influenza Virus mRNAs[b]:		
NS (A strain)	CAG | GTAGAG	$\begin{matrix}T\\C\end{matrix}$ $_9$CCAG | G
NS (B strain)	GAG | GTGGGT	GATCGGACAG | T
M (A strain):		
M2	AAC | GTATGT	$\begin{matrix}T\\C\end{matrix}$ $_7$G(A)$_4$(T)$_3$GCAG | G
mRNA$_3$	CAG | GTAGAT	

[a] From reference 70.
[b] The influenza viral mRNA junction sequences are from references 14, 57, 60.
[c] The two sequences underlined with a solid line, GT and AG, are found at all splice junctions in eukaryotic mRNAs. The two influenza virus nucleotides underlined with a dotted line are found only very rarely at the splice junctions of other eukaryotic mRNAs.

ments). Consequently, when these influenza viral sequences are transcribed from a DNA template by RNA polymerase II, they are recogniced and utilized by cellular splicing enzymes. This strongly suggests, but does not prove, that cellular enzymes in the nucleus splice influenza viral mRNAs when they are transcribed from an RNA template (vRNA) by the viral transcriptase. In this regard, it should be noted that the pattern of interrupted and uninterrupted mRNAs observed when the M and NS DNAs are expressed in the SV40 vector differs in several respects from that observed in influenza virus-infected cells. First, M2 mRNA is not made in the SV40 system [59], indicating that the M2 donor site (which contains an unusual nucleotide—a C residue—at the 3′ end of the leader or 5′ exon [Table 1]) is not used. These investigators suggest that failure to use this donor site may somehow be caused by the 5′ or 3′ SV40 sequences present in the M1 mRNA precursor. It will be of interest to determine whether M2 mRNA is produced when much of these 5′ and 3′ SV40 sequences are removed, or when the $mRNA_3$ donor site is removed, or when the M DNA is expressed in a different vector. Also, the relative abundances of the steady state amounts of interrupted and uninterrupted mRNAs in the SV40 system differ from those in influenza virus-infected cells. In the SV40 system, the amounts of the interrupted mRNA ($mRNA_3$ or NS2 mRNA) and of the uninterrupted mRNA (M1 or NS1 mRNA) produced were similar [59], whereas in influenza virus-infected cells the amount of the interrupted mRNA(s) was only at most 5% of that of the respective uninterrupted mRNA [55, 60]. This difference has also been attributed to some effect of the SV40 5′ and 3′ sequences in the M1 mRNA precursor on the efficiency of splicing [59]. It is not known how the production of interrupted viral mRNAs is controlled at a low level in influenza virus-infected cells. Nor is it known why the production of NS2 mRNA (and possibly the other interrupted viral mRNAs) apparently requires early protein synthesis [56]. One possibility is that a fraction of the nuclear M1 and NS1 viral mRNAs needs to be modified in some way by a virus-coded protein to make these mRNAs suitable substrates for splicing.

Cellular enzymes in the nucleus have also been implicated in the other major modification of influenza viral mRNAs, the methylation of their internal A residues. In the total viral mRNA population, there are approximately two m^6A residues per chain in the virus-coded region (*i.e.*, after the host-donated primer sequences have been removed) [43]. No activity capable of methylating the internal A residues of the viral mRNAs has so far been detected in purified virions (Beaton, Plotch, and Krug, unpublished experiments), consistent with this activity being host-coded. If this is the case, the m^6A residues in the virus-coded region of the viral mRNAs would be expected to be in the same sequences (GAC and AAC) as in DNA-directed RNA polymerase II transcripts [16, 23, 87, 101, 102]. Recent experiments in the author's laboratory indicate that the m^6A residues in the virus-coded region of the viral mRNAs are indeed in GAC and AAC sequences (Beaton and Krug, unpublished experiments).

The role of internal m^6A residues is not known, but it has been postulated that these residues, which have been shown to occur at a specific subset of the available GAC and AAC sequences in several DNA-directed transcripts [2, 7, 16], constitute part of the recognition sites for splicing or at least influence the splicing reaction. Several lines of evidence are consistent with such a role for internal m^6A residues:

(i) internal methylation of the initial RNA transcripts from the adenovirus late promoter occurs soon after they are synthesized and before they are spliced [18]; (ii) the m^6A-containing sequences in the initial transcripts from the late adenovirus transcripts are completely or nearly completely conserved in the stable mRNAs in the cytoplasm [18]; and (iii) inhibition of internal methylation (and also 2′-O methylation of the 5′ penultimate base) of retrovirus RNA by cycloleucine inhibits splicing of this RNA [96]. If m^6A residues serve a role in the splicing of influenza viral mRNAs, the simplest situation would be that only viral mRNAs that undergo splicing (the M1 and NS1 mRNAs) would contain m^6A residues in the virus-coded region. This is not the case: recent experiments in the author's laboratory indicate that all eight of the major viral mRNAs contain m^6A in the virus-coded region (Beaton and Krug, unpublished experiments). It is not known whether the M1 and NS1 mRNAs contain more m^6A residues per chain than the other viral mRNAs nor whether some or all of the m^6A residues in the M1 and NS1 mRNAs are strategically located with respect to the putative splice junctions.

Evidence has also been obtained suggesting that the host-coded primer sequences also contain an m^6A residue. ^3H-methyl-labeled *in vivo* viral mRNA was purified so that all of the radiolabel was apparently in viral mRNA molecules, and this RNA preparation was hybridized to vRNA and digested with pancreatic of T1 RNase: along with the 5′ methylated cap structure, one of the three m^6A residues found per viral mRNA chain was not protected [43]. To rule out the possibility that the unprotected m^6A is actually in intact host cell RNAs present in low amounts in the viral mRNA preparation, it will be necessary to verify this result using newer methods of viral mRNA purification, *i.e.*, selection with cloned viral DNAs. If verified, it will be important to establish whether the m^6A in the host-donated primer region is in GAC or AAC, the sequences methylated by host enzymes. In the cloned DNAs containing a copy of the host-donated primer, neither GAC or AAC sequences have been found [42], but as discussed previously these host-donated sequences may not be representative of the overall population. The function of the m^6A residue apparently present in the primer sequence is not known. Such a residue is not required for priming, since mRNAs lacking m^6A (*e.g.*, A1MV RNA 4) are effective primers *in vitro* [42]. Nonetheless, m^6A could still be recognized by the viral endonuclease and/or initiating protein, thereby increasing the efficiency of priming. The m^6A in the host-coded region could already be present in the cellular capped RNAs used as primers; alternatively, methylation could occur during or after transfer of the host primer sequences to the viral mRNA molecules. If the former were the case, the viral enzymes would be using a specific subset of cellular transcripts that contain an m^6A residue within the first 10 to 13 nucleotides at the 5′ end.

III. Synthesis of Full-Length Transcripts

Full-length complementary copies of the vRNA segments were first identified in experiments in which ^3H-uridine-labeled infected cell RNA was hybridized to vRNA, and the resulting hybrids were treated with S1 nuclease and analyzed by gel

electrophoresis [29]. It was noted that the hybrids containing unpolyadenylylated (poly [A] [—]) transcripts migrated slower than the corresponding hybrids containing the polyadenylated transcripts (*i.e.*, the viral mRNAs), and that the migration of the hybrids containing poly (A) (—) transcripts was not changed by the S1 nuclease treatment. These results suggested that the poly (A) (—) transcripts are complete copies of the vRNA segments. Several lines of evidence have verified this conclusion [28, 31]. In fact, direct sequencing of the 3′ ends of the poly (A) (—) transcripts showed that the 13 nucleotides at the 3′ ends are complementary to the 13-nucleotide sequence found at the 5′ ends of each of the eight vRNA segments [32]. Consequently, the termination of transcription which occurs during viral mRNA synthesis at a site 17 to 22 nucleotides from the 5′ ends of the vRNA templates does not operate during the synthesis of the poly (A) (—) full-length transcripts. In addition, the evidence indicates that initiation of the full-length transcripts occurs without a primer at the 3′-terminal U of the vRNA (32) (instead of at the 3′-penultimate C as occurs during viral mRNA synthesis [77]). These transcripts have a pppA 5′-terminus, and their 5′ sequences are complementary to the 3′-terminal sequences of the vRNA templates [32].

Because they are complete copies of the vRNA segments (and lack host-derived primer sequences), these transcripts most likely function as the templates for vRNA synthesis (replication). Consistent with this role, the full-length transcripts are not associated with polyribosomes, but instead are apparently in the form of nucleocapsids [29]. Because the complements of the 5′ ends of the vRNA segments are present in the full-length transcripts, the 5′ and 3′ ends of these transcripts have the potential of forming duplex "panhandle" structures similar to those formed by the vRNAs. It has been postulated that this duplex structure serves as a recognition signal for the polymerase which synthesizes both vRNA and the full-length transcripts, thereby precluding the viral mRNAs from being copied [28]. The present author, however, considers it more likely that these duplex "panhandle" structures serve as a recognition signal for the nucleocapsid protein to form specific nucleocapsid structures and that only those RNAs, *i.e.*, full-length transcripts and vRNA, that form such specific nucleocapsid structures can associate with, and consequently be copied by, polymerase proteins (most likely P proteins).

Unlike the viral mRNAs, the synthesis of the full-length transcripts is not regulated: approximately equimolar amounts of each of the eight full-length transcripts are synthesized throughout infection [29]. The full-length transcripts comprise only about 5% of the viral RNA transcripts synthesized in the infected cell. Their synthesis is inhibited by cycloheximide [4, 29], strongly suggesting that the synthesis of one or more proteins, presumably virus-specified, is required in order to modify the transcriptase complex so that it initiates transcription without a primer and continues transcription past the termination site utilized during viral mRNA synthesis.

The full-length transcripts are probably synthesized in the nucleus, where the viral mRNAs are also synthesized [33]. The evidence for this conclusion comes from the same experiment that showed that the eight viral mRNAs are synthesized in the nucleus (Fig. 5). In this experiment, the ^{3}H-uridine-pulse-labeled RNA in the nuclear and cytoplasmic fractions of infected cells was separately annealed to excess

vRNA, and the resulting hybrids after RNase T2 digestion were analyzed by gel electrophoresis. ^{3}H-labeled hybrids with the mobility of full-length transcripts were detected predominantly in the nuclear sample (arrows in Fig. 5). It is not known whether these full-length transcripts remain in the nucleus to direct vRNA synthesis, or whether, like the viral mRNAs, they are first transported to the cytoplasm.

The mechanism of the switch from mRNA to full-length transcript synthesis is not known. Because capped RNA primers are not used to initiate the synthesis of full-length transcripts, the viral PB2 protein, the cap-recognizing protein, would not be expected to participate in this initiation. One possibility is that the PB1 protein, which catalyzes the initiation of transcription during viral mRNA synthesis, is somehow modified so that it can also initiate transcription of full-length transcripts, *i.e.*, initiate transcription without a primer and at the terminal U rather than at the penultimate C of the vRNA templates. Such a modification of the PB1 protein could result from the action of the virus-coded protein(s) whose synthesis is required for full-length synthesis. The identity of this putative virus-coded protein(s) is not known. The available evidence argues against this protein(s) being the NS protein(s): two *ts* virus mutants with a defect in the NS genome segment are not defective in full-length synthesis at the nonpermissive temperature [104]. The other change from viral mRNA synthesis (the absence of transcriptional termination at the site 17 to 22 nucelotides from the 5′ ends of the vRNA templates) suggests that the duplex „panhandle" structures formed between the 5′ and 3′ ends of the vRNA templates are opened up during the synthesis of the full-length transcripts. How this might occur is not known.

IV. Synthesis of vRNA (Replication)

The data concerning influenza vRNA synthesis are meager and in some respects conflicting. Different time courses of vRNA synthesis have been described by different groups of investigators. Results have been obtained indicating that: (i) maximal rates of vRNA synthesis occur at early times, by about 2.5 hours of infection, essentially coincident with the maximal rates of viral mRNA synthesis [32]; or (ii) most vRNA synthesis occurs after the peak of viral mRNA synthesis [4, 67, 89, 98]. Hopefully, this discrepancy will be resolved when better methods of measuring vRNA synthesis become available. One group of investigators has found that vRNA synthesis is regulated, *i.e.*, different relative proportions of the eight vRNAs are produced at various times after infection [32]. At early times, the NP and NS vRNAs were preferentially synthesized. Subsequently, the rates of synthesis of the other vRNAs increased, and the rate of synthesis of the NS vRNA decreased relative to that of the M (and other) vRNAs. Because this pattern of vRNA synthesis is similar to the relative rates of synthesis of the different viral mRNAs during infection, it was proposed that the regulation of viral mRNA synthesis occurs at the level of vRNA synthesis (see section entitled *Regulation of viral mRNA Synthesis in the Infected Cell*). The only exceptions to this similarity between vRNA and viral mRNA synthesis were with the three P segments: the rate of synthesis of

the P vRNAs was always about 10-fold greater than the rate of synthesis of the P mRNAs. Because the putative templates for vRNA synthesis—the poly (A) (−) full-length transcripts—are made in approximately equimolar amounts throughout infection [29, 32], it was concluded that selective copying of different templates occurs at various times after infection.

Not much else is known about vRNA synthesis. Temperature-shift experiments conducted *in vivo* with *ts* virus mutants have implicated at least two of the P proteins in vRNA synthesis [62], but the precise role of these P proteins and possibly of other virus-coded proteins in catalyzing vRNA synthesis is not known. The cellular site (nucleus or cytoplasm) of vRNA synthesis is not known. In addition, it is not known how approximately equimolar amounts of the different vRNA segments are packaged into virions.

References

[1] Allen, H., McCauley, J., Waterfield, M., Gething, M. J.: Influenza virus RNA segment 7 has the coding capacity for two polypeptides. Virology *107,* 548–551 (1980).

[2] Aloni, Y., Dhar, R., Khoury, G.: Methylation of nuclear simian virus 40 RNAs. J. Virol. *32,* 52–60 (1979).

[3] Bannerjee, A. K.: 5′-terminal cap structure in eukaryotic messenger ribonucleic acids. Microbiol. Rev. *44,* 175–205 (1980).

[4] Barrett, T., Wolstenholme, A. J., Mahy, B. W. J.: Transcription and replication of influenza virus RNA. Virology *98,* 211–225 (1979).

[5] Barry, R. D., Ives, D. R., Cruickshank, J. G.: Participation of deoxyribonucleic acid in the multiplication of influenza virus. Nature *194,* 1139–1140 (1962).

[6] Beaton, A. R., Krug, R. M.: Selected host cell capped RNA fragments prime influenza viral RNA transcription *in vivo.* Nucl. Acids Res. *9,* 4423–4436 (1981).

[7] Beemon, K., Keith, J.: Localization of N^6-methyladenosine in the Rous sarcoma virus genome. J. Mol. Biol. *113,* 165–179 (1977).

[8] Blaas, D., Patzelt, E., Kuechler, E.: Cap-recognizing protein of influenza virus. Virology *116,* 339–348 (1982).

[9] Blaas, D., Patzelt, E., Keuchler, E.: Identification of the cap binding protein of influenza virus. Nucl. Acids Res. *10,* 4803–4812 (1982).

[10] Bouloy, M., Morgan, M. A., Shatkin, A. J., Krug, R. M.: Cap and internal nucleotides of reovirus mRNA primers are incorporated into influenza viral complementary RNA during transcription *in vitro.* J. Virol. *32,* 895–904 (1979).

[11] Bouloy, M., Plotch, S. J., Krug, R. M.: Globin mRNAs are primers for the transcription of influenza viral RNA *in vitro.* Proc. Natl. Acad. Sci. U.S.A. *75,* 4886–4890 (1978).

[12] Bouloy, M., Plotch, S. J., Krug, R. M.: Both the 7-methyl and 2′-O-methyl groups in the cap of a mRNA strongly influence its ability to act as a primer for influenza viral RNA transcriptions. Proc. Natl. Acad. Sci. U.S.A. *77,* 3952–3956 (1980).

[13] Briedis, D. J., Conti, G., Munn, E. A., Mahy, B W. J.: Migration of influenza virus-specific polypeptides from cytoplasm to nucleus of infected cells. Virology *111,* 154–164 (1981).

[14] Briedis, D. J., Lamb, R. A.: Influenza B virus genome: sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS1 and NS2 proteins. J. Virol. *42,* 186–193 (1982).

[15] Caliguiri, L. A., Compans, R. W.: Analysis of the *in vitro* product of an RNA-dependent RNA polymerase isolated from influenza virus-infected cells. J. Virol. *14,* 191–197 (1974).

[16] Canaani, D., Kahana, C., Lavi, S., Groner, Y.: Identification and mapping of N^6-methyladenosine containing sequences in Simian Virus 40 RNA. Nucl. Acids Res. *6,* 2879–2899 (1979).

[17] Caton, A. J., Robertson, J. S.: Stucture of the host-derived sequences present at the 5′ ends of influenza virus mRNA. Nucl. Acids Res. *8,* 2591–2603 (1980).
[18] Chen-Kiang, S., Nevins, J. R., Darnell, J. E.: N-6-Methyladenosine in adenovirus type 2 nuclear RNA is conserved in the formation of messenger RNA. J. Mol. Biol. *135,* 733–752 (1979).
[19] Compans, R. W., Caliguiri, L. A.: Isolation and properties of an RNA polymerase from influenza virus-infected cells. J. Virol. *11,* 441–448 (1973).
[20] Compans, R. W., Content, J., Duesberg, P. H.: Structure of the ribonucleoprotein of influenza virus. J. Virol. *10,* 795–800 (1972).
[21] Davis, A. R., Hiti, A. L., Nayak, D. P.: Influenza defective interfering viral RNA is formed by internal deletion of genomic RNA. Proc. Natl. Acad. Sci. U.S.A. *77,* 215–219 (1980).
[22] Dhar, R., Chanock, R. M., Lai, C.-J.: Nonviral oligonucleotides at the 5′ terminus of cytoplasmic influenza viral mRNA deduced from cloned complete genomic sequences. Cell *21,* 495–500 (1980).
[23] Dimock, K., Stoltzfus, C. M.: Sequence specificity of internal methylation in B77 avian sarcoma virus RNA subunits. Biochemistry *16,* 471–478 (1977).
[24] Etkind, P. R., Krug, R. M.: Influenza viral messenger RNA. Virology *62,* 38–45 (1974).
[25] Etkind, P. R., Krug, R. M.: Purification of influenza viral complementary RNA: its genetic content and activity in wheat germ cell-free extracts. J. Virol. *16,* 1464–1475 (1975).
[26] Gupta, K. C., Kingsbury, D. W.: Conserved polyadenylation signals in two negative-strand RNA virus families. Virology *120,* 518–523 (1982).
[27] Hastie, N. D., Mahy, B. W. J.: RNA-dependent RNA polymerase in nuclei of cells infected with influenza virus. J. Virol. *12,* 951–961 (1973).
[28] Hay, A. J., Abraham, G., Skehel, J. J., Smith, J. C., Fellner, P.: Influenza virus messenger RNAs are incomplete transcripts of the genome RNAs. Nucl. Acids Res. *4,* 4179–4209 (1977).
[29] Hay, A. J., Lomniczi, B., Bellamy, A. R., Skehel, J. J.: Transcription of the influenza virus genome. Virology *83,* 337–355 (1977).
[30] Hay, A. J., Skehel, J. J.: Studies on the synthesis of influenza virus proteins. In: Negative Strand Viruses (Mahy, B. W. J., Barry, R. D., eds.), Vol. 2, 635–655. New York-London: Academic Press 1975.
[31] Hay, A. J., Skehel, J. J., McCauley, J.: Structure and synthesis of influenza virus complementary RNAs. Phil. Trans. R. Soc. (Lond.) *B 288,* 341–348 (1980).
[32] Hay, A. J., Skehel, J. J., McCauley, J.: Characterization of influenza virus RNA complete transcripts. Virology *116,* 517–522 (1982).
[33] Herz, C., Stavnezer, E., Krug, R. M., Gurney, T., jr.: Influenza virus, an RNA virus, synthesizes its messenger RNA in the nucleus of infected cells. Cell *26,* 391–400 (1981).
[34] Horisberger, M. A.: The large P proteins of influenza A viruses are composed of one acidic and two basic polypeptides. Virology *107,* 302–305 (1980).
[35] Inglis, S. C., Barrett, T., Brown, C. M., Almond, J. W.: The smallest genome RNA segment of influenza virus contains two genes that may overlap. Proc. Natl. Acad. Sci. U.S.A. *76,* 3790–3794 (1979).
[36] Inglis, S. C., Carroll, A. R., Lamb, R. A., Mahy, B. W. J.: Polypeptides specified by the influenza virus genome. I. Evidence for eight distinct gene products specified by fowl plague virus. Virology *74,* 489–503 (1976).
[37] Inglis, S. C., Mahy, B. W. J.: Polypeptides specified by the influenza virus genome. 3. Control of synthesis in infected cells. Virology *95,* 154–164 (1979).
[38] Jackson, D. A., Caton, A. J., McCready, S. J., Cook, P. R.: Influenza virus RNA is synthesized at fixed sites in the nucleus. Nature *296,* 366–368 (1982).
[39] Koennecke, I., Boschek, C. B., Scholtissek, C.: Isolation and properties of a temperature-sensitive mutant (*ts* 412) of an influenza A virus recombinant with a *ts* lesion in the gene coding for the nonstructural protein. Virology *110,* 16–25 (1981).
[40] Krug, R. M.: Influenza viral RNPs newly synthesized during the latent period of viral growth in MDCK cells. Virology *44,* 125–136 (1971).
[41] Krug, R. M.: Cytoplasmic and nucleoplasmic viral RNPs in influenza virus-infected MDCK cells. Virology *50,* 103–113 (1972).
[42] Krug, R. M.: Priming of influenza viral RNA transcription by capped heterologous RNAs. Current Topics in Microbiol. and Immunol. *93,* 125–150 (1981).

[43] Krug, R. M., Broni, B. A., Bouloy, M.: Are the 5′ ends of influenza viral mRNAs synthesized *in vivo* donated by host mRNAs? Cell *18,* 329–334 (1979).
[44] Krug, R. M., Broni, B. A., LaFiandra, A. J., Morgan, M. A., Shatkin, A. J.: Priming and inhibitory activities of RNAs for the influenza viral transcriptase do not require base-pairing with the virion RNA template. Proc. Natl. Acad. Sci. U.S.A. *77,* 5874–5878 (1980).
[45] Krug, R. M., Etkind, P. E.: Cytoplasmic and nuclear virus-specific proteins in influenza virus-infected MDCK cells. Virology *56,* 334–348 (1973).
[46] Krug, R. M., Morgan, M. M., Shatkin, A. J.: Influenza viral messenger RNA contains internal N^6-methyladenosine and 5′-terminal 7-methylguanosine in cap structures. J. Virol. *20,* 45–53 (1976).
[47] Krug, R. M., Plotch, S. J., Ulmanen, I., Herz, C., Bouloy, M.: The mechanism of initiation of influenza viral RNA transcription. In: The Replication of Negative-Strand Viruses (Compans, R. W., Bishop, D. H. L., eds.), 291–302. Elsevier/North-Holland 1981.
[48] Krug, R. M., Soeiro, R.: Studies on the intranuclear localization of influenza virus-specific proteins. Virology *64,* 378–387 (1975).
[49] Krug, R. M., Ueda, M., Palese, P.: Temperature-sensitive mutants of influenza WSN virus defective in virus-specific RNA synthesis. J. Virol. *16,* 790–796 (1975).
[50] Lamb, R. A., Briedis, D. J., Lai, C.-J., Choppin, P. W.: Multiple mRNAs and coding regions derived from individual influenza A and B virus RNA segments. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., ed.), 141–158. Academic Press 1981.
[51] Lamb, R. A., Choppin, P. W.: Synthesis of influenza virus protein in infected cells: translation of viral proteins including three P polypeptides, from RNA produced by primary transcription. Virology *74,* 504–519 (1976).
[52] Lamb, R. A., Choppin, P. W.: Synthesis of influenza virus polypeptides in cells resistant to alpha-amanitin: evidence for the involvement of cellular RNA polymerase II in virus replication. J. Virol. *23,* 816–819 (1977).
[53] Lamb, R. A., Choppin, P. W.: Segment 8 of the Influenza virus genome is unique in coding for two polypeptides. Proc. Natl. Acad. Sci. U.S.A. *76,* 4908–4912 (1979).
[54] Lamb, R. A., Choppin, P. W.: Identification of a second protein (M2) encoded by RNA segment 7 of influenza virus. Virol. *112,* 729–737 (1981).
[55] Lamb, R. A., Choppin, P. W., Chanock, R. M., Lai, C.-J.: Mapping of the two overlapping genes for polypeptides NS1 and NS2 on RNA segment 8 of influenza virus genome. Proc. Natl. Acad. Sci. U.S.A. *77,* 1857–1861 (1980).
[56] Lamb, R. A., Etkind, P. R., Choppin, P. W.: Evidence for a ninth influenza virus polypeptide. Virology *91,* 60–78 (1978).
[57] Lamb, R. A., Lai, C.-J.: Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. Cell *21,* 475–485 (1980).
[58] Lamb, R. A., Lai, C.-J.: Conservation of the influenza virus membrane protein (M1) amino acid sequence and an open reading frame of RNA segment 7 encoding a second protein (M2) in H1N1 and H3N2 strains. Virology *112,* 746–751 (1981).
[59] Lamb, R. A., Lai, C.-J.: Spliced and unspliced messenger RNAs synthesized from cloned influenza virus M DNA in an SV40 vector: Expression of the influenza virus membrane protein (M1). Virology *123,* 237–256 (1982).
[60] Lamb, R. A., Lai, C.-J., Choppin, P. W.: Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: Colinear and interrupted mRNAs code for overlapping proteins. Proc. Natl. Acad. Sci. U.S.A. *78,* 4170–4174 (1981).
[61] Lazarowitz, S. G., Compans, R. W., Choppin, P. W.: Influenza virus structural and non-structural proteins in infected cells and their plasma membranes. Virology *46,* 830–843 (1971).
[62] Mahy, B. W. J., Barrett, T., Nichol, S. T., Penn, C. R., Wostenholme, A. J.: Analysis of the functions of influenza virus genome RNA segments by use of temperature-sensitive mutants of fowl plague virus. In: The Replication of Negative Strand Viruses–Developments in Cell Biology, Vol. 7 (Bishop, D. H. L., Compans, R. W., eds.), 379–387. Elsevier/North-Holland. 1981.
[63] Mahy, B. W. J., Bromley, P. A.: *In vitro* product of a ribonucleic acid polymerase induced by influenza virus. J. Virol. *6,* 259–268 (1970).

[64] Mahy, B. W. J., Hastie, N. D., Armstrong, S. J.: Inhibition of influenza virus replication by α-amanitin: mode of action. Proc. Natl. Acad. Sci. U.S.A. *69,* 1421–1424 (1972).
[65] Mahy, B. W. J., Hastie, N. D., Raper, R. H., Brownson, J. M. T., Carroll, A. R.: RNA polymerase activities of nuclei from influenza virus-infected cells. In: Negative Strand Viruses (Mahy, B W. J., Barry, R. D., eds.), 445–467. New York: Academic Press 1975.
[66] Manley, J. L., Sharp, P. A., Gefter, M. L.: RNA synthesis in isolated nuclei. Processing of adenovirus serotype 2 late messenger RNA precursors. J. Mol. Biol. *159,* 581–599 (1982).
[67] Mark, G. E., Taylor, J. M., Broni, B. B., Krug, R. M.: Nuclear accumulation of influenza viral RNA transcripts and the effects of cycloheximide, actinomycin D, and α-amanitin. J. Virol. *29,* 744–752 (1979).
[68] McGeoch, D., Kitron, N.: Influenza virion RNA-dependent RNA polymerase: stimulation by guanosine and related compounds. J. Virol. *15,* 686–695 (1975).
[69] Meier-Ewert, H., Compans, R. W.: Time course of synthesis and assembly of influenza virus proteins. J. Virol. *14,* 1083–1091 (1974).
[70] Mount, S. M.: A catalog of splice junction sequences. Nucl. Acids Res. *10,* 459–472 (1982).
[71] Mowshowitz, S. L.: P1 is required for initiation of cRNA synthesis in WSN influenza virus. Virology *91,* 493–495 (1978).
[72] Mowshowitz, S. L., Ueda, M.: Temperature-sensitive virion transcriptase activity in mutants of WSN influenza virus. Arch. Virol. *52,* 135–141 (1976).
[73] Nevins, J. R., Darnell, J. E., jr.: Steps in the processing of Ad2 mRNA: poly A^+ nuclear sequences are conserved and poly A addition precedes splicing. Cell *15,* 1477–1493 (1978).
[74] Nichol, T., Penn, C. R., Mahy, B. W. J.: Evidence for the involvement of influenza A (fowl plague Rostock) virus protein P2 in ApG and mRNA primed *in vitro* RNA synthesis. J. Gen. Virol. *57,* 407–413 (1981).
[75] Palese, P.: The genes of influenza virus. Cell *10,* 1–10 (1977).
[76] Plotch, S. J., Bouloy, M., Krug, R. M.: Transfer of 5′ terminal cap of globin mRNA to influenza viral complementary RNA during transcription *in vitro*. Proc. Natl. Acad. Sci. U.S.A. *76,* 1618–1622 (1979).
[77] Plotch, S. J., Bouloy, M., Ulmanen, I., Krug, R. M.: A unique cap (m^7GppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. Cell *23,* 847–858 (1981).
[78] Plotch, S. J., Krug, R. M.: Influenza virion transcriptase: the synthesis *in vitro* of large, polyadenylic acid-containing complementary RNA. J. Virol. *21,* 24–34 (1977).
[79] Plotch, S. J., Krug, R. M.: Segments of influenza virus complementary RNA synthesized *in vitro*. J. Virol. *25,* 579–586 (1978).
[80] Plotch, S. J., Tomasz, J., Krug, R. M.: Absence of detectable capping and methylating enzymes in influenza virions, J. Virol. *28,* 75–83 (1978).
[81] Revel, M., Groner, Y.: Post-transcriptional and translational controls of gene expression in eukaryotes. Ann. Rev. of Biochem. *47,* 1079–1126 (1978).
[82] Robertson, H. D., Dickson, E., Plotch, S. J., Krug, R. M.: Identification of the RNA region transferred from a representative primer, β-globin mRNA, to influenza mRNA during *in vitro* transcription. Nucl. Acids Res. *8,* 925–942 (1980).
[83] Robertson, J. S.: 5′ and 3′ terminal nucleotide sequences of the RNA genome segments of influenza virus. Nucl. Acids Res. *6,* 3745–3757 (1979).
[84] Robertson, J. S., Schubert, M., Lazzarini, R. A.: Polyadenylation sites for influenza virus mRNA. J. Virol. *38,* 157–163 (1981).
[85] Rott, R., Scholtissek, C.: Specific inhibition of influenza replication by a α-amanitin. Nature *228,* 56 (1970).
[86] Salditt-Geogrieff, M., Harpold, M., Chen-Kiang, S., Darnell, J. E., jr.: The addition of 5′ cap structures occurs early in hnRNA synthesis and prematurely terminated molecules are capped. Cell *19,* 69–78 (1981).
[87] Schibler, U., Kelly, D. E., Perry, R. P.: Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. J. Mol. Biol. *115,* 695–714 (1977).
[88] Scholtissek, C., Rott, R.: Ribonucleic acid nucleotidyl tranferase induced in chick fibroblasts after infection with an influenza virus. J. Gen. Virol. *4,* 125–137 (1969).

[89] Scholtissek, C., Rott, R.: Synthesis *in vivo* of influenza virus plus and minus strand RNA and its preferential inhibition by antibiotics. Virology *40,* 989–996 (1970).
[90] Schubert, M., Keene, J. D., Herman, R. C., Lazzarini, R. A.: Site on the vesicular stomatitis virus genome specifying polyadenylation and the end of the L gene mRNA. J. Virol. *34,* 550–559 (1980).
[91] Shatkin, A. J.: Capping of eukaryotic mRNAs. Cell *9,* 645–653 (1976).
[92] Skehel, J. J., Burke, D. C.: Ribonucleic acid synthesis in chick embryo cells infected with fowl plaque virus. J. Virol. *3,* 429–438 (1969).
[93] Skehel, J. J., Hay, A. J.: Nucleotide sequences at the 5′ termini of influenza virus RNAs and their transcripts. Nucl. Acids Res. *5,* 1207–1219 (1978).
[94] Smith, G. L., Hay, A. J.: Replication of the influenza virus genome. Virology *118,* 96–108 (1982).
[95] Spooner, L. L. R., Barry R. D.: Participation of DNA-dependent RNA polymerase II in replication of influenza viruses. Nature *268,* 650–652 (1977).
[96] Stoltzfus, C. M., Dane R. W.: Accumulation of spliced avian retrovirus mRNA is inhibited in S-adenosylmethionine-depleted chicken embryo fibroblasts. J. Virol. *42,* 918–931 (1982).
[97] Taylor, J. M., Hampson, A. W., White, D. O.: The polypeptides of influenza virus. I. Cytoplasmic synthesis and nuclear accumulation. Virology *39,* 419–425 (1969).
[98] Taylor, J. M., Illmensee, R., Litwin, S., Herring, L., Broni, B. A., Krug, R. M.: The use of specific radioactive probes to study the transcription and replication of the influenza virus genome. J. Virol. *21,* 530–540 (1977).
[99] Ulmanen, I., Broni, B. A., Krug, R. M.: The role of two of the influenza virus core P proteins in recognizing cap 1 structures (m^7GpppNm) on RNAs and in initiation viral RNA transcription. Proc. Natl. Acad. Sci. U.S.A. *78,* 7355–7359 (1981).
[100] Ulmanen, I., Broni, B. A., Krug, R. M.: Influenza virus temperature-sensitive cap (m^7GpppNm)-dependent endonuclease. J. Virol. *45,* 27–35 (1983).
[101] Wei, C-M., Gershowitz, A., Moss, B.: 5′-terminal and internal methylated nucleotide sequences in HeLa cell mRNA. Biochemistry *15,* 397–401 (1976).
[102] Wei, C-M., Moss, B.: Nucleotide sequence at the N^6-methyladenosine sites of HeLa cell messenger ribonucleic acid. Biochemistry *16,* 1672–1676 (1977).
[103] Winter, G., Fields, S.: Cloning of influenza cDNA into M13: the sequence of the RNA segment encoding the A/PR/8/34 matrix protein. Nucl. Acids Res. *8,* 1965–1974 (1980).
[104] Wolstenholme, A. J., Barrett, T., Nichol, S. T., Mahy, B. W. J.: Influenza virus-specific RNA and protein syntheses in cells infected with temperature-sensitive mutants defective in the genome segment encoding nonstructural proteins. J. Virol. *35,* 1–7 (1980).

4

Genetic Relatedness of Influenza Viruses (RNA and Protein)

C. Scholtissek

I. Introduction

The genomes of influenza A and B viruses are comprised of 8 single-stranded RNA segments of different molecular weights, while influenza C viruses contain only 7 segments (for reviews see Palese, 1977; Scholtissek, 1978, 1979 a; Palese *et al.*, 1980). This unusual structure of the influenza virus genome makes a genetic comparison of different isolates interesting, because during double infection of a host with two different strains of the same type, each RNA segment behaves like a chromosome, and free reassortment is possible, creating new strains. Thus, when comparing two strains, each pair of allelic RNA segments has to be studied separately.

Various analytical methods have been developed and applied to study the genetic relatedness of influenza viruses. Each method has its specific value and limitation. Some are relatively simple and can be used for screening a large number of new isolates. Others are more tedious and time- or material-consuming, but give more definitive answers. In this chapter, the various techniques will be described and their advantages and disadvantages will be discussed. Since most of the comparative data concerning genetic relatedness are available for influenza A viruses, I will concentrate on the various subtypes of these viruses.

II. Genetic Relatedness of Viral RNAs

A. Differences in Migration Rates of the RNA Segments As Revealed by Polyacrylamide Gel Electrophoresis (PAGE)

The RNA segments of influenza viruses (vRNA) can be labeled *in vivo* or *in vitro* and they can be separated by polyacrylamide gel electrophoresis in the presence of 6 M urea according to differences in size and secondary structure. The migration rates of allelic genes (*i.e.* RNA segments) of different subtype strains or of strains belonging to the same subtype can vary significantly.

Analysis of gel electrophoresis patterns of viral RNA molecules has greatly aided in providing answers to the following problems:

i) Differences in migration rates of the RNA segments on PAGE under limiting denaturing conditions are specific for various prototype strains. They can be correlated with differences in migration rates of the corresponding viral proteins. By constructing specific recombinants, especially using temperature-sensitive mutants, assignment of the various RNA segments to corresponding viral proteins and/or functions can be achieved (for reviews see Palese, 1977; Scholtissek, 1978, 1979 a; Palese and Schulman, 1976; Ritchey *et al.*, 1976; Palese *et al.*, 1977 a; Harms *et al.*, 1978; Almond and Barry, 1979; Heller and Scholtissek, 1980). Such an assignment of the RNA segments of the avian influenza A virus strain A/fowl plague/Ros-

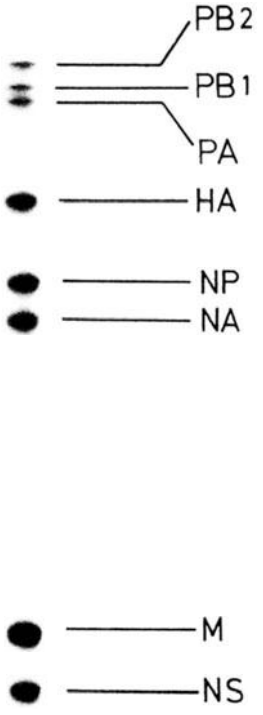


Fig. 1. Polyacrylamide gel electrophoresis of A/fowl plague/Rostock/34 (H7N1) virus ^{32}P-labeled virus was purified, and its RNA was isolated. The RNA segments were separated by electrophoresis on a polyacrylamide slab gel in the presence of 6 M urea. The RNA is visualized by exposure of an X-ray film to the gel (Scholtissek *et al.*, 1976). The RNA segments are numbered according to their migration rates, segment 1 (PB2) being the slowest moving one. PB2 (= Pol 1 = P3), cap recognizing protein; PB1 (= Ptra = P1), initiating protein; PA (= Pol 2 = P2), elongating protein (Horisberger, 1980, 1982; Krug, 1982). The assignments of the P3 and P1 protein are at variance to those shown by Scholtissek (1978, 1979). The migration rates of vRNA segments 1 and 2 of PR8, which have been used formerly for comparative assignments seem to differ in different laboratories due to slight changes of the separation conditions, *e.g.* the concentration of the urea in the electrophoresis buffer. HA, hemagglutinin; NP, nucleoprotein; NA, neuraminidase; M, matrix or membrane protein; NS, nonstructural protein (by courtesy of V. von Hoyningen)

tock/34 (FPV, H7N1) to the corresponding gene products is shown in Fig. 1. This method is now widely used for the determination of the derivation of RNA segments in recombinants made in the laboratory.

ii) By comparing the migration rates of the RNA segments of parental and progeny virions, it has been demonstrated that genetic reassortment between influenza A viruses occurs in the intestinal tract of naturally- and laboratory-infected ducks (Hinshaw *et al.*, 1980).

iii) Comparison of RNA migration patterns in polyacrylamide gels in the presence of urea has been used as an epidemiological tool for the identification of independent influenza A virus isolates (Palese and Ritchey, 1977 b; Hinshaw *et al.*, 1978). Differences in electrophoretic migration rates of RNA segments of recent influenza B virus isolates have also been recognized (Hugentobler *et al.*, 1981). However, it should be noted that a few point mutations may alter the secondary structure of an RNA and may significantly change its electrophoretic migration rate in polyacrylamide gels (Scholtissek *et al.*, 1978 b).

If the secondary structure of RNAs is largely eliminated by treatment with glyoxal, the P, NP, M, and NS genes of one virus strain show migration patterns in PAGE which are indistinguishable from those of other strains. Under these denaturing conditions, only the RNA segments coding for the glycoproteins, the hemagglutinin (HA) and the neuraminidase (NA), exhibit significant differences in migration rates indicating real differences in molecular weights between virus strains (Desselberger and Palese, 1978).

B. Molecular Hybridization

1. Direct RNA-RNA Hybridization

Double-stranded RNA is relatively stable against treatment with pancreatic RNase or T1-RNase in the presence of 2 × SSC (0.3 M NaCl, 0.03 M Na-citrate), while single-stranded RNA is completely digested to acid-soluble oligonucleotides. The influenza virus-specific RNA polymerase obtained from infected cells synthesizes *in vitro* an RNA which is complementary to vRNA (Scholtissek, 1969). This cRNA can be radioactively labeled if the synthesis proceeds in the presence of labeled nucleoside triphosphates. The product then can be converted to double-stranded RNA by hybridization with a surplus of non-labeled vRNA of the same strain. If vRNA of another strain is used for hybridization, the double-stranded structure of that RNA exhibits considerable mismatching. Such mismatched regions are not resistant to RNase and are digested to acid-soluble material. Thus, by hybridization and treatment with RNase one obtains a relative measure of base sequence homology between two genetically related strains. By this technique, it was found that various influenza A virus strains differ significantly in the base sequences of their RNAs (Scholtissek and Rott, 1969).

When the vRNAs of different influenza virus strains are labeled by growing the virus in the presence of ^{32}P-orthophosphate, individual RNA segments can be isolated following PAGE. These labeled segments are then used for hybridization with an excess of non-labeled cRNA isolated from the microsomal fraction of

infected chick embryo cells. By applying this technique, the 8 RNA segments of different influenza A viruses can be divided into two groups: those which are highly conserved, and those which are much less conserved when different subtypes are compared. The vRNA segments coding for the two glycoproteins, HA and NA, are less conserved, while the segments coding for the 3 P proteins, the NP and M proteins are more conserved (Scholtissek *et al.*, 1976, 1977 a; Rohde *et al.*, 1977; Harms *et al.*, 1978). With respect to the base sequence of the smallest vRNA segment, coding for the NS-proteins, the influenza A viruses can be divided clearly into two groups (Scholtissek and von Hoyningen-Huene, 1980).

The direct RNA-RNA hybridization technique has been very useful in the establishment of a new nomenclature for influenza A viruses (WHO-Bulletin, 1980). The RNase protection data suggest that the hemagglutinins of strains previously classified as H0, H1 and Hsw1 viruses all belong to the same subtype (now H1) (Scholtissek *et al.*, 1977 a). Similarly, the hemagglutinins of the former H3, Heq2, and Hav7 subtypes are now combined into the H3 subtype. The hemagglutinin gene of the A/duck/Ukraine/1/63 (Hav7 Neq2) strain was found to exhibit the highest RNase protection (92%) when it was hybridized to the RNA of the human A/Hong Kong/1/68 (H3N2) virus. It is interesting to note that the A/duck/Ukraine/1/63 strain was isolated before the appearance of the first H3 prototype in the human population (Scholtissek *et al.*, 1978 a). The hemagglutinin genes of later Hav7 isolates from birds (A/heron/Chabarovsk/700/73; A/duck/Chabarovsk/1610/72; A/duck/Chabarovsk/698/73) showed an RNase protection of 100% following hybridization to the RNA of the H3 Hong Kong strain (Scholtissek, unpublished). In these cases, however, it cannot be excluded that the original human H3 virus HA found its way back into an animal reservoir. However, it has to be kept in mind that the 3 isolates from Chabarovsk are not identical to the human H3 strains, since their RNA segments 8 exhibit an RNase protection of only about 40% when hybridized to those of human influenza A virus strains (Scholtissek and von Hoyningen-Huene, 1980). This latter finding suggests the reassortment of genes derived from human and animal strains. The RNase protection of hybrid molecules between RNA segments 4 of Heq1 and Hav1 strains was greater than 60%, indicating that these strains also form one subtype (now H7) (Scholtissek, unpublished).

With respect to the neuraminidase genes, it has been suggested on the basis of serological data that the Nav2, Nav3, and Nav6 subtypes constitute one group (WHO-Bulletin, 1979). The hybridization data revealed that the NA genes of the Nav2 and Nav3 strains tested (now N3) are indeed highly related (80% and greater RNase protection), while protection of the NA genes of Nav2 and Nav3 strains with the RNA of Nav6 strains was only about 20% (Scholtissek *et al.*, 1980). This suggests that the NA genes of the latter strains represent a separate subgroup (now N9).

Different degrees of RNase protection between members of different HA subtypes have been discovered, too. For example, while RNase protection between the HA genes of most of the subtypes tested was in the order of 30%, that between a member of H3 (Hong Kong), and a member of the H7 (fowl plague) strain was 43% (Scholtissek *et al.*, 1976). This is in accordance with the evolutionary tree of the various influenza A virus hemagglutinins, as revealed by direct RNA sequencing (see below, Fig. 3).

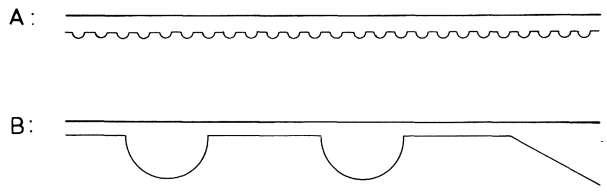


Fig. 2. Alternative schemes for the distribution of 50% mismatching at random *(A)* or in 3 clusters *(B)*. See text for discussion

The direct hybridization technique has been useful for identification of the derivation of various RNA segments in artificial recombinants constructed *in vitro* (Scholtissek *et al.*, 1977). Similarly, this method was successfully employed for the identification of the new 1977 pandemic strain (A/USSR/90/77; H1N1), which was found to be genetically almost identical to isolates from 1950 (A/Fort Warren/1/50, H1N1) (Scholtissek *et al.*, 1978). In addition, this technique aided in determining genetic drift among different influenza virus strains (Scholtissek, 1980).

During a comparative study of the HA genes of various pathogenic and non-pathogenic avian H7 strains it turned out that several strains with a cleavable HA were relatively distantly related genetically, while some H7 strains with a cleavable HA were closely related to the HA genes of strains with a noncleavable HA. This suggests that the evolution of the majority of sequences within H7 HA genes is independent of the evolution of the cleavage site (Bosch *et al.*, 1982).

The limitations of the direct hybridization method are as follows: (1) If differences in base sequences of two RNAs become extensive and widely dispersed within the molecules, hybrids are unstable and may be digested by RNase. For example, if there is only 50% base sequence homology between two RNAs, the distribution of changes may occur as outlined in Fig. 2: In the upper part, "A", mismatching is evenly distributed over the total length of the molecule. In this case, RNase protection would be negligible. In the lower part, "B", the homologous sequences are clustered in a few totally conserved domains. In this case, RNase protection would be indeed 50%. A situation similar to the latter model is true for the RNA segments coding for HA and NA: The residual RNase protection of hybrid molecules between subtypes is about 30% and 20%, respectively, and this residual protection is due to highly conserved domains, as was demonstrated by melting profiles. Thus, mutations leading to amino acid replacements seem to be lethal in these domains, while in the variable domains many mutations may be tolerated (Scholtissek, 1979 b). In most instances, when RNase protection of RNA hybrids is relatively low, the technique yields values which underestimate the base sequence homology of the RNAs (see below). (2) Single point mutations cannot be detected by this hybridization technique. Even if an RNase cleaves a heterologous double-stranded RNA molecule at a position of a point mutation, the material still would be acid insoluble. However, the method can be rendered quite sensitive, and can be used for the discrimination of highly releated genes, if prior to RNase treatment the hybrid molecules are heated in the presence of 1% formaldehyde. In this case, mismatching due to point mutations causes a nucleation point of melting and single

stranded regions are formed which are fixed by formaldehyde. The method has been calibrated by introducing point mutations at random using nitrous acid and measuring melting profiles of the hybrid RNAs in relation to the number of mutations. Thus, 10 point mutations in an RNA molecule of the length of the HA gene would be recognized (Scholtissek, 1980).

2. *Competitive Hybridization*

For competitive hybridization, homologous double-stranded RNA molecules are prepared by annealing ^{125}I-labeled vRNA segments (Bean *et al.*, 1980 a) to unlabeled cRNA. The labeled double-stranded RNA is then melted and allowed to reanneal in the presence of increasing amounts of an unlabeled allelic vRNA segment of a genetically related strain. If the unlabeled vRNA contains an RNA sequence identical or partly homologous to that of the ^{125}I-labeled vRNA the unlabeled vRNA will compete for annealing sites on the cRNA and will prevent completely or partially the reannealing of the labeled vRNA. This labeled vRNA then becomes digestable with RNase. If RNase resistance is plotted against the amount of different competing RNAs, the resulting plateaus permit identification of the genetic relatedness of influenza virus genes (Bean *et al.*, 1980 b). The limitations of this method are similar to those described for direct RNA-RNA hybridization (Sriram *et al.*, 1980). The advantage is that very little material is needed for hybridization. However, the number of the samples needed to reach a plateau is relatively high, especially if two allelic RNA segments are distantly related (Bean *et al.*, 1980 b).

Using this method, it was found that there is a greater genetic diversity among cocirculating avian Hav7 Neq2 (H3N8) strains of a geographic area compared to the diversity found among human H3N2 strains isolated 9 years apart (Sriram *et al.*, 1980). Furthermore, naturally occurring recombinants between recent cocirculating H1N1 and H3N2 human pandemic strains were detected employing this procedure (Bean *et al.*, 1980 b). An influenza A virus isolated from harbor seals has been characterized genetically by the competitive hybridization technique. All RNA segments of this virus were more closely related to different avian influenza virus strains than to other mammalian virus strains. This finding may mean that new mammalian viruses might be created by reassortment of avian influenza virus strains (Bean *et al.*, 1981; Webster *et al.*, 1981).

3. *DNA-RNA Hybridization*

A relatively stable ^{3}H-labeled DNA probe for hybridization can be prepared by transcribing individual vRNA segments using reverse transcriptase in the presence of actinomycin D and ^{3}H-dNTP. The single-stranded complementary DNA (cDNA) can then be hybridized with unlabeled vRNA of homologous or heterologous strains. In the heterologous hybridization, single-stranded regions are digested by nuclease S1 (Palese *et al.*, 1981). Using this method, it was found that the genomes of several type A strains exhibited a higher variation than those of type B strains, and that the genomes of B strains varied more than those of C strains (Palese *et al.*, 1981). Furthermore, the genetic relatedness of segment 8 (NS-gene) of various

strains has been studied. As found by the direct RNA-RNA hybridization technique, at least two different groups of influenza A virus NS genes were recognized (Palese, personal communication).

The homology values obtained by this method are somewhat lower than those found using the direct RNA-RNA hybridization technique. However, the advantage of the DNA-RNA hybridization technique is that once a batch of labeled DNA has been prepared, the probes can be used over a long period of time.

4. *Analysis of Double-Stranded Nuclease S1-Treated Hybrid Molecules by Polyacrylamide Gel Electrophoresis*

This technique takes advantage of the fact that after infection in the presence of cycloheximide, viral cRNA is synthesized (Scholtissek and Rott, 1970) which can be labeled with ^{3}H-uridine (Hay *et al.*, 1977). This labeled cRNA then may be hybridized to an excess of unlabeled homologous or heterologous vRNA. After treatment with nuclease S1, the double-stranded molecules are analyzed by polyacrylamide gel electrophoresis. Extensive nuclease S1 digestion destroys all single-stranded RNA and leaves only the homologous double-stranded RNA segments intact. This technique is simple and can be used for the analysis of recombinants derived *in vitro* from different subtype strains (Hay *et al.*, 1977; Ghendon *et al.*, 1979, 1981 a). Limited nuclease S1 treatment removes only the single-stranded tails from the hybrid molecules and prevents exclusion of the partially double-stranded molecules by the gel. However, this treatment does not destroy heterologous hybrids of genetically highly-related strains, but such hybrid molecules appear to migrate more slowly in PAGE than the corresponding homologous RNA-RNA hybrids. This technique has been used to compare human influenza viruses which are closely related chronologically. Differences were detected between the H3N2 isolates, A/Port Chalmers/73 and A/Scotland/74, in RNA segments 1, 2, 4, 6, and 7 (Hay *et al.*, 1977). Similar studies have been performed with other highly related viruses isolated during epidemics (Ghendon *et al.*, 1981 b; Klimov and Ghendon, 1981).

The advantage of this method is that a stable ^{3}H-cRNA probe is used and that it is sensitive and relatively simple. The disadvantage of the procedure is that it does not provide any quantitative data. Although the technique is sensitive, it certainly does not allow detection of point mutations in an RNA. Furthermore, the sensitivity of the method may interfere with the proper identification of RNA segments in recombinants in which one of the parents has been partially inactivated or treated with mutagens, since silent mutations might be picked up by this technique (Hay *et al.*, 1977).

C. Oligonucleotide Fingerprints

RNA can be digested by endonucleases which cleave the RNA after a specific base, leading to a mixture of nucleotides and oligonucleotides of different lenghts. The cleavage products are separable by two-dimensional gel electrophoresis (De

Wachter and Fiers, 1972) or by homochromatography after electrophoresis. If the viral RNA is labeled *in vivo* by growing virus in the presence of ^{32}P, large quantities of ^{32}P-phosphate have to be used. A more sensitive method is to digest unlabeled RNA first and to label the oligonucleotides at their 5′ termini using γ-^{32}P-ATP and polynucleotide kinase prior to two-dimensional gel electrophoresis (Frisby, 1977). Since T1 ribonuclease cleaves after guanine residues, it yields 3′-GMP and oligonucleotides ending with G. The large oligonucleotides are of interest, since they may be unique for a particular RNA. When the large oligonucleotides of two genetically related strains are compared, only a percentage of each genome is visualized. Thus, in such a comparison many point mutations must escape detection. Furthermore, oligonucleotides with the same lengths and base compositions, but different sequences, are indistinguishable in most separation systems. Another limitation of oligonucleotide map analysis is that if overall base sequence homology of two RNAs is less than 90%, the oligonucleotide maps will share few, if any, unique oligonucleotides (Aaronson *et al.*, 1982).

The oligonucleotide fingerprint technique was applied to show that the 8 RNA segments of influenza A viruses are unique and not breakdown products of a large molecule (McGeoch *et al.*, 1976). Furthermore, it was shown that the M protein genes of two strains of fowl plague virus differed in three unique oligonucleotides (Hay *et al.*, 1977). Similarly, the genetic variability of the Hong Kong virus has been studied with this technique (Ortin *et al.*, 1980). By oligonucleotide mapping, it was demonstrated that all RNA segments of H1N1 isolates causing a pandemic in 1977 were very similar to those of 1950 strains, suggesting that a 1950 strain had remained dormant over a period of many years (Nakajima *et al.*, 1978). A more recent oligonucleotide fingerprint study by Kozlov *et al.* (1981) suggests that 7 RNA segments of the A/USSR/90/77 strain were highly related to the A/Fort Warren/1/50 isolate. Only the M-protein genes (RNA segment 7) of these two strains were less related. Another study using oligonucleotide map analysis revealed that some viruses of the H1N1 subtype isolated since 1977 were recombinants between viruses of the cocirculating H1N1 and H3N2 subtypes (Young and Palese, 1979).

Oligonucleotide fingerprint analysis and partial sequencing of the RNAs of human H1N1 strains isolated in 1977 suggested divergent evolution of these strains from a common ancestor. Such an ancestor strain was shown to be related to strains circulating around 1950 (Young *et al.*, 1979).

The genomes of duck influenza A viruses isolated at a lake in Northern France in 1976 have been analyzed by the same technique. Of the two strains compared, only the genes coding for the hemagglutinin and the M protein were highly related, while the other genes exhibited significantly greater variation. This is taken as further evidence that recombination of animal influenza viruses occurs under natural conditions (Desselberger *et al.*, 1978).

A comparative oligonucleotide fingerprint analysis of the RNAs of recently isolated human and swine influenza A viruses demonstrated that the RNAs of these viruses were very similar. This finding suggested that swine H3N2 strains isolated in East Asia were derived from contemporary human strains (Nerome *et al.*, 1981).

The results of an analysis of the late H2N2 and early H3N2 strains isolated in 1967/68 suggested that an H2N2 virus circulating during or shortly before the first

appearance of the H3N2 strain has contributed seven genes to H3N2 virus by reassortment (Nakajima *et al.*, 1982).

A corresponding study of human influenza C viruses isolated between 1949 and 1979 revealed marked genetic stability over a long period of time (Meier-Ewert *et al.*, 1981).

D. Sequencing of RNA Segments

1. Sequencing of ^{32}P-End Labeled RNAs by Partial Nuclease Digestion

The first comparative sequences of the 5′ and 3′ termini of influenza A and B virus RNAs were obtained by the sequencing technique developed by Donis-Keller *et al.* (1977). Viral RNA was treated with alkaline phosphatase to remove the 5′-terminal phosphate. Subsequently, the RNA was incubated with ^{32}P-ATP and polynucleotide kinase. Aliquots of end-labeled RNAs were partially digested with RNases of different cleavage specificities, and the resulting labeled oligonucleotide mixtures were separated by polyacrylamide gel electrophoresis. Using this technique, it was found that at the 5′ termini of fowl plague virus RNAs the sequence of the first 13 nucleotides was the same for all 8 segments. Furthermore, a cluster of 5 to 7 uridine residues was observed at positions 17–23. At the 3′ end, the sequence of the first 12 nucleotides was also shared by all 8 segments. When different strains were compared, it was found that the sequences of all the terminal nucleotides were highly conserved, including those of influenza B and C virus RNAs (Skehel and Hay, 1978; Robertson, 1979; Palese, 1980; Desselberger *et al.*, 1980).

The technique described has not been widely applied for the sequencing of influenza virus RNA segments, since easier and more reliable methods have recently been developed.

2. Sequencing of the 3′ End of RNAs Using the Dideoxy Method

The dideoxy method developed originally by Sanger *et al.* (1977) has been adapted for sequencing RNA segments of influenza viruses (Air, 1979; Hall and Air, 1981). For this purpose, viral RNA segments were polyadenylated by the riboadenylate transferase of *E. coli.* Then, using p(dT8dA) as primer, virus-specific cDNA was synthesized. Alternatively, a synthetic doclecamer specific for the common 3′-terminal sequence was used to prime, avoiding the addition of poly (A). The reaction mixture also contained one of the four 2′,3′-dideoxynucleoside triphosphates. Under these conditions, occasionally a 2′,3′-dideoxynucleotide is incorporated in place of the corresponding deoxynucleotide, so that chain termination occurs. The mixtures of labeled oligodeoxynucleotides were separated by polyacrylamide gel electrophoresis. The resulting pattern then permitted derivation of the nucleotide sequence.

Comparisons of the 3′-terminal sequences of 220 bases of segment 7 (matrix [M] protein gene) and of 230 bases of segment 8 (nonstructural [NS] protein gene) of human influenza virus strains isolated between 1934 and 1977 suggested that these segments were conserved during the evolution of H3N2 from H1N1. Thirteen

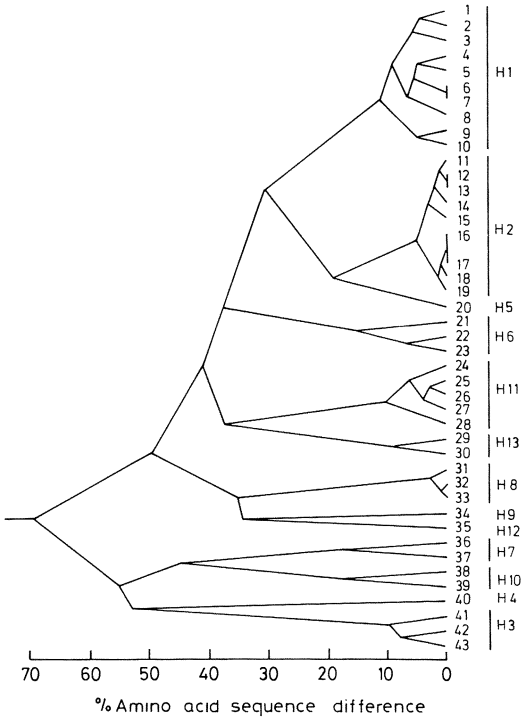


Fig. 3. Dendrogram showing the relationships between the HA1 N-terminal amino acid sequences deduced from genomic sequences of viruses representing the 13 known subtypes (taken from Webster *et al.*, 1982)

The sequences were aligned using amino acids that are invariable in certain positions of all sequences (amino acid position from the N terminus of most subtypes: Cys 4, Gly 6, Thr 18, Val 26, Cys 42, Cys 55, Gly 63, Pro 65, Cys 67, Glu 81). The percentage sequence differences of all pairwise comparisons of the aligned sequences starting from the N-terminal Asp or corresponding amino acid were used to calculate the dendrogram. Thus, the positions of each bifurcation in the dendrogram indicate the mean sequence difference of the sequences connected through that point. The sequences used in the analysis are: 1, A/NWS/33 (H1N1); 2, A/PR/8/34 (H1N1); 3, A/BH/35 (H1N1); 4, A/Bel/42 (H1N1); 5, A/Loyang/4/51 (H1N1); 6, A/USSR/90/77 (H1N1); 7, A/Fort Warren/1/50 (H1N1); 8, A/duck/Alberta/35/76 (H1N1); 9, A/swine/Wisconsin/15/30 (H1N1); 10, A/New Jersey/11/76 (H1N1); 11, A/RI/57 (H2N2); 12, A/Tokyo/3/67 (H2N2); 13, A/Berkeley/68 (H2N2); 14, A/Netherlands/68 (H2N2); 15, A/duck/GDR/72 (H2N9); 16, A/duck/New York/6750/78 (H2N2); 17, A/duck/Ontario/77 (H2N1); 18, A/duck/Alberta/77/77 (H2N3); 19, A/pintail/Alberta/293/77 (H2N9); 20, A/shearwater/Australia/75 (H5N3); 21, A/shearwater/Australia/72 (H6N5); 22, A/duck/Alberta/1298/79 (H6N5); 23, A/duck/Alberta/1297/79 (H6N9); 24, A/tern/Australia/75 (H11N9); 25, A/duck/Ukraine/1/60 (H11N9); 26, A/duck/England/56 (H11N6); 27, A/duck/New York/12/78 (H11N6); 28, A/duck/Memphis/576/76 (H11N9); 29, A/gull/Mass/26/80 (H13N6); 30, A/gull/Md/704/77 (H13N6); 31, A/turkey/Ontario/6118/68 (H8N4); 32, A/duck/Alberta/827/78 (H8N4); 33, A/duck/Alberta/283/77 (H8N4); 34, A/turkey/Wisconsin/1/66 (H9N2); 35, A/duck/Alberta/60/76 (H12N5); 36, A/turkey/Oregon/71 (H7N3); 37, A/equine/Prague/1/56 (H7N7); 38, A/duck/Manitoba/53 (H10N7); 39, A/quail/Italy/1117/65 (H10N8); 40, A/duck/Alberta/28/76 (H4N6); 41, A/black duck/Australia/702/78 (H3N8); 42, A/Memphis/1/71 (H3N2); 43, A/duck/Ukraine/1/63 (H3N8)

base replacements were found within the sequence of 220 bases of the M genes when the 1934 and 1977 isolates were compared, and 22 replacements were observed within the first 230 bases of the NS genes. However, most of these replacements did not lead to changes in the sequence of amino acids. Thus, although these genes are highly conserved, genetic drift can be recognized (Hall and Air, 1981). No significant base sequence homology was found between RNA segments 8 of influenza A and B viruses (Air and Hackett, 1980).

A thorough study of sequence relationships among the hemagglutinin genes of 43 influenza A virus strains from the 13 known subtypes has been performed by Air and colleagues (Air, 1981; Webster *et al.*, 1981; Hinshaw *et al.*, 1982). This was done by determining sequences of about 300 nucleotides at the 3′ end of RNA segment 4, coding for HA. By calculating the "dissimilarities" in deduced amino acid sequences, a dendrogram was constructed, which is shown in Fig. 3. The H1 and H2 hemagglutinins are highly related, which was found also by determining the total sequence of these genes (Hiti *et al.*, 1981, see below). However, the dissimilarities among the various subtypes are greater than those found in genes within a subtype. Some amino acids, especially cysteine residues, are completely conserved among the hemagglutinins of the different subtypes, indicating that these genes have evolved from a single ancestor (Air, 1981).

A similar study was performed on the neuraminidase gene by Blok and Air (1980). The sequences of the 200 3′-terminal nucleotides of six N1 and five N2 influenza A virus genes were compared. Within subtypes, the base replacements were about 10%, while between subtypes only about 100 out of 200 nucleotides were found to be identical. Only at the N terminus, a highly conserved domain of 11 to 12 amino acids was recognized, as deduced from the nucleotide sequence. Some of the N1 subtypes exhibited block deletions of 33 to 48 nucleotides. This correlates with differences in the mobilities of RNA segments 6 from different subtypes in PAGE (Blok and Air, 1982).

The advantage of the dideoxy method as it is used here is that the genes of many strains can be easily compared. The disadvantage is that only about 20% of each gene is analyzed. An extrapolation of the data to the entire molecule may not be valid for genes in which highly conserved domains are followed by variable domains. This criticism is especially relevant for genes in which block deletions were observed (Blok and Air, 1982).

3. Sequencing of Total RNA Segments

The most precise information on genetic relatedness of different influenza virus genes will be obtained by comparing the total base sequences of allelic RNA segments. However, the determination of the total sequences of different RNA molecules is still quite tedious and time consuming.

The general procedures for obtaining complete RNA sequences are as follows: Total vRNA or individual vRNA segments are polyadenylated at their 3′ ends and are reverse-transcribed with an oligo (dT) primer. Thereafter, full-length double-stranded DNA (dsDNA) copies are purified by polyacrylamide gel electrophoresis. The dsDNA molecules are treated with S1 nuclease and ligated (after appropriate

modifications at their ends) into a suitable plasmid. After transformation of *E. coli*, clones with plasmids containing full-length DNA inserts are selected, and the influenza virus-specific DNA is excised by restriction enzymes. Each restriction fragment is then sequenced by the method of Maxam and Gilbert (1977). This standard technique has been widely used for sequencing influenza virus genes (Emtage *et al.*, 1979; Porter *et al.*, 1979; Min Jou *et al.*, 1980). It has been modified and improved by using the specific dodecamer primer, which is complementary to the 3′ end of the RNA segments, already discussed (Baez *et al.*, 1980; Hiti *et al.*, 1981). A further improvement is the use of a second synthetic primer complementary to the 3′ end of the first strand cDNA for the synthesis of the second DNA strand (Baez *et al.*, 1980; Winter *et al.*, 1981 a). Sequencing of influenza virus genes can be expedited by combination of the latter technique with the "shot-gun" approach: dsDNA from the influenza virus RNA segments is digested with restriction enzymes and the resulting fragments are ligated to the replicative form (RF) of bacteriophage M13 mp2 DNA. Single-stranded template DNA is prepared from recombinant phages and sequenced by the dideoxy technique using a primer flanking the site of insertion (Winter and Fields, 1980). With the help of these techniques, the total sequence of many RNA segments of different influenza virus strains has been determined (see chapter 2).

a) Hemagglutinin Gene

The first complete sequence of an influenza virus RNA segment was obtained for the hemagglutinin (HA) gene of fowl plague virus (A/FPV/Rostock/34) (Porter *et al.*, 1979). Since then, several HA genes belonging to 4 different subtypes have been sequenced (Min Jou *et al.*, 1980; Verhoeyen *et al.*, 1980; Gething *et al.*, 1980; Both and Sleigh, 1980; Hiti *et al.*, 1981; Winter *et al.*, 1981 b; Fang *et al.*, 1981; Both and Sleigh, 1981; Sleigh *et al.*, 1981). The sequences can be directly compared by aligning conserved cysteine, tryptophan and proline residues of the deduced proteins. In this way deletions and/or insertions can be recognized.

The deduced amino acid sequences of different HA molecules reveal the presence of hydrophobic signal peptides 15 to 18 amino acids in length, which are highly variable among the subtypes. Even within subtypes, variation is exceptionally great within this region. In addition, there are two other regions of hydrophobic amino acids in the HAs: One is located at the amino terminus of the HA_2 subunit and is highly conserved for the first 14 amino acids. The other, which is less conserved among subtypes, is close to the carboxyl terminus of the HA_2 subunit.

Although the cysteine residues are highly conserved in the HA genes of all subtypes, the positions of the carbohydrate side chains vary among different strains. With respect to amino acid replacements in general, there are domains in the HA genes which are highly conserved among subtypes. Variation within the HA_1 subunit is higher than in the HA_2 subunit. Some subtypes are more closely related to each other than others. For example, the HA genes of an H1 and an H2 subtype exhibit a conservation of deduced amino acids and nucleotides (in parenthesis) of 67% (65%), whereas the conservation between H1 and H3 hemagglutinins is only 42% (50%). The H3 hemagglutinin again is more closely related to an H7 hemagglutinin (49% [50%]) when compared with an H1 hemagglutinin (42%

[50%]) (Gething *et al.*, 1980; Hiti *et al.*, 1981). These data are consistent with the dendrogram in Fig. 3. However, there is clearly less variation between members of the same subtype compared with other subtypes. Thus, a clear distinction between genetic (and antigenic) shift and drift can be made (Gething *et al.*, 1980).

When the sequences of HA genes belonging to the same subtype are compared, base substitutions leading to amino acid replacements accumulate according to the chronologic order of isolation of strains. When the sites of replacement are located within the three-dimensional structure of the hemagglutinin molecule (Wilson *et al.*, 1981; Wiley *et al.*, 1981), they appear to cluster in 4 different domains. These sites are thought to represent the biologically relevant antigenic areas of the HA molecules at the surface of the virion (Both *et al.*, 1980; Both and Sleigh, 1980; Verhoeyen *et al.*, 1980; Wiley *et al.*, 1981; Both and Sleigh, 1981; Sleigh *et al.*, 1981; Fang *et al.*, 1981; Webster *et al.*, 1982). Under the selective pressure of the host immune system, mutants which carry alterations within the antigenic sites of the surface glycoproteins may be selected leading to changes in the immunological properties of the virus (antigenic drift).

It had already been shown by other methods that the shift in 1957 from the H1N1 to the H2N2 subtype was accompanied by a replacement of 4 RNA segments in the new subtype. Four RNA segments in the H2N2 subtype viruses were derived from viruses circulating in the human population at that time. With respect to the H3N2 subtype, only the HA gene was replaced, while the other 7 RNA segments were retained from the H2N2 subtype (Scholtissek *et al.*, 1978; Nakajima *et al.*, 1982). The question then arose, where did the new RNA segments come from? In the latter case, it has been suggested on the basis of serological (Laver and Webster, 1973) as well as hybridization data (Scholtissek *et al.*, 1978) that the HA gene of a duck influenza virus (A/duck/Ukraine/63, H3N8) might be the closest ancestor of the new human H3N2 strain. Comparison of the HA gene of the duck Ukraine virus with that of the human "Aichi" strain from 1968, and of the A/Victoria/3/75 strain, confirmed this close relationship. The duck isolate is more closely related to the Aichi strain than the Aichi virus to the Victoria strain, which was derived from the latter by drift. Based on the accumulation of silent base changes in these genes, a common progenitor hemagglutinin was postulated. It was speculated that such a progenitor virus circulated around 1950 in an animal reservoir (Fang *et al.*, 1981).

b) Neuraminidase Gene

For the neuraminidase (NA), which is the second surface glycoprotein of influenza viruses, only three complete sequences are available (Fields *et al.*, 1981; Hiti *et al.*, 1982; Markoff and Lai, 1982). The NA RNAs of A/PR/8/34 (H1N1) and A/WSN/34 (H1N1) viruses code for polypeptides of 453 and 454 amino acids, respectively. The RNA of the H3N2 strain (A/Udorn/72) directs the synthesis of a polypeptide of 469 amino acids.

65 base changes and 32 amino acid changes were found in the NA genes of the PR8 and WSN strains. In addition, there is a stretch of 6 amino acids starting at position 58 from the N terminus, which is completely different between these two strains. Another fundamental difference is that the WSN strain contains only three

potential glycosylation sites compared with 5 encoded in the PR8 NA gene. It is not yet known which of these differences in the NA molecules is responsible for differences in neurovirulence and host range. The PR8 strain and a progenitor (WS) of the WSN strain were isolated in 1934 and 1933, respectively. Unfortunately, these strains have been passaged so many times in different laboratories under different conditions, that many of the observed changes might have occurred in the laboratory. Consequently, the original relatedness of these two NA genes might have been much higher. When compared with RNA segments 1 of these two strains (see below), the variation in amino acid replacements of the NA genes is much greater, possibly reflecting a greater ability of NA function to survive mutation.

There is very little sequence homology between the NA genes of the N1 and N2 subtypes, in agreement with the hybridization data (Scholtissek *et al.*, 1976). However, regions of homology were detected if the cysteine residues of the deduced PR8 and Udorn NA proteins were aligned. At the N terminus there is a conserved sequence of 12 amino acids followed by 79 amino acids with almost no homology, including a deletion (or insertion) of 15 amino acids. Relatively short homologies are scattered over the rest of the molecule. From amino acid positions 7 to 34 there is a hydrophobic domain, by which the NA molecule is presumably inserted into the lipid bilayer of the virus envelope (see also Blok *et al.*, 1982). The low homology between the N1 and N2 neuraminidase genes compared with the high homology between several corresponding genes of H1N1 and H2N2 strains (see below) suggests that the antigenic shift in 1957 most likely was the result of reassortment between an animal and a human influenza virus strain. This finding confirms earlier data based on hybridization experiments (Scholtissek *et al.*, 1978).

c) The Three P-Protein Genes

The complete sequences of RNA segments 1, 2, and 3 coding for the 3 P-proteins of PR8 virus are known. There is no significant homology among these sequences (Fields and Winter, 1982; Winter and Fields, 1982). RNA segment 3 of other strains has not yet been sequenced. However, the total sequence is known for RNA segment 1 of another human influenza A virus strain (A/WSN/33, H1N1) (Kaptein and Nayak, 1982). As mentioned above, the PR8 and WSN strains belong to the same subtype and were isolated at about the same time. Therefore, they might be expected to be genetically highly related. Both code for PB2 polypeptides which are 759 amino acids long. There are several consensus donor and acceptor sites for splicing in these sequences. Consequently, it is not yet known whether additional smaller mRNAs are coded by RNA segment 1. According to the deduced amino acid sequence, the PB2 protein coded by RNA segment 1 is a basic protein. Between the sequences of the two strains, 22 amino acid changes were found. As discussed above, these differences may be due in part to the different passage histories of the two strains.

RNA segments 2 of A/PR/8/34 (H1N1) and A/NT/60/68 (H3N2) viruses were sequenced and analyzed. Both RNAs code for the PB1 proteins. The A/NT/60/68 (H3N2) strain (Bishop *et al.*, 1982) was isolated 34 years after the PR8 strain, and, therefore, significant differences were expected. Among many silent mutations

there are only 21 amino acid replacements scattered over the total length of the molecule, which codes for a polypeptide of 757 amino acids. From nucleotide residues 583 to 669, there are 15 basic amino acids. Although there are three amino acid replacements in this region, there is no change in net charge.

d) The Nucleoprotein Gene

The sequences of two nucleoprotein (NP) genes belonging to two different subtype strains are known (A/PR/8/34; H1N1, and A/NT/60/68; H3N2). The sequence of segment 5 (NP gene) of A/PR/8/34 has been determined by two different groups (Van Rompuy *et al.*, 1981; Winter and Fields, 1981). The discrepancy between the two published structures presumably is due to the fact that Van Rompuy *et al.*, (1981) missed a small Hinf 1 fragment during sequencing. The sequence has been corrected in the meanwhile (Van Rompuy *et al.*, 1982). For comparison, the data obtained by Winter and Fields (1981) will be used. The number of nucleotides (1565) and deduced amino acids (498) of the nucleoprotein genes of the two subtypes is the same. The proteins are highly basic, as expected for a protein which binds to nucleic acids. There were only 29 amino acid replacements, despite the fact that the two viruses were isolated 34 years apart, suggesting that the NP gene has been carried over during the two antigenic shifts (Winter and Fields, 1981; Huddleston and Brownlee, 1982).

e) The Membrane Protein Gene

The sequences of the membrane (M) protein genes of two human (PR8, H1N1; and A/Udorn/72, H3N2) and of one avian (A/FPV/Rostock/34, H7N1) influenza A viruses have been determined (Winter and Fields, 1980; Allen *et al.*, 1980; Lamb and Lai, 1981; McCauley *et al.*, 1982). There are two open reading frames in the cRNA of all three strains which lead—if they are accordingly translated—in all three cases to two proteins of 252 (MW 27,000) and 96 (MW 11,000) amino acids, respectively. Corresponding polypeptides have been isolated from infected cells: the membrane (M1) protein and another nonstructural protein (M2), the function of which is not yet known (Lamb and Choppin, 1981). The structure of the M gene is discussed in detail in the chapter by R. A. Lamb.

Silent mutations which do not lead to amino acid changes are scattered over the total length of RNA segment 7, except for the region of overlapping reading frames (68 nucleotides), indicating that there are no hot spots of mutation. The sequence of the M protein gene is highly conserved when compared with those of other RNA segments, especially with those of genes coding for the surface glycoproteins. Comparing the PR8 (H1N1) M gene with that of the Udorn (H3N2) strain, there were only 7 amino acids out of 252 replaced within the M1 protein. Between the M2 proteins, 11 amino acids out of 96 were changed. This indicates that the M2 proteins are less conserved than the M1 proteins. A comparison with the FPV M protein gene suggests that the PR8 M protein gene is more closely related to the former than to the Udorn M protein gene.

There are several regions of 10 or more uncharged residues. Especially from residues 115 to 151 there are only two charged amino acids, but 16 hydrophobic

amino acids. It is possible to construct a three-dimensional structure of the M1 protein with a highly hydrophobic domain, which might be embedded in the lipid membrane, and a more hydrophilic domain interacting possibly with the nucleocapsid. No gross amino acid replacements were found within these domains.

One other observation deserves attention: The M protein gene of PR8 has been sequenced by two different groups (Winter and Fields, 1980; Allen *et al.*, 1980). Six nucleotide differences were found between these sequences, one leading to the replacement of an amino acid. These differences might either represent changes in the viral genome during passage of the virus in different laboratories introduced after the original isolation in 1934, or might be due to sequencing errors.

Recently, the sequence of segment 7 of an influenza B virus (B/Lee/40) has become available (Briedis *et al.*, 1982). There are also two open reading frames. One may code for a protein (M1) of 248 amino acids; the other may code for a polypeptide of 195 amino acids in a +2 reading frame.

When the M1 protein of the B/Lee strain is compared with that of the A/Udorn/72 strain, base sequence homologies can be detected. Out of 248 amino acids, 63 are conserved, including an internal hydrophobic domain with three cysteine residues at corresponding positions. This observation confirms comparative hybridization data between the M genes of an avian influenza A virus (FPV) and another influenza B virus, although the hybridization data are somewhat higher than expected (Scholtissek *et al.*, 1977 b). No significant amino acid homology has been detected in the deduced amino acid sequences of the M2 proteins of the B/Lee and A/Udorn viruses.

f) The Nonstructural Protein Gene

The total nucleotide sequence of RNA segment 8 (NS gene) has been determined for 4 different influenza A virus subtypes (Porter *et al.*, 1980; Lamb and Lay, 1980; Baez *et al.*, 1980; Winter *et al.*, 1981; Baez *et al.*, 1981) and for one influenza B virus (Briedis and Lamb, 1982). Three of the influenza A virus subtypes have highly conserved NS genes in agreement with hybridization data (Scholtissek *et al.*, 1976; Scholtissek and von Hoyningen-Huene, 1980), while the NS gene of the A/Duck/Alberta/60/76 virus (H12N5) exhibits a high degree of variation. This finding is in agreement with direct RNA-RNA hybridization data (Scholtissek, unpublished).

On all vRNA segments 8, including that of an influenza B virus, two different types of mRNAs are synthesized; one consists of the normal length with a stop codon at a site translating about 80% of the mRNA leading to the NS1 protein. The other is derived by splicing, continuing in a +1 reading frame, and leading to the NS2 protein. The structure of the NS gene has been discussed in more detail in the chapter by R. A. Lamb.

Silent mutations are scattered over the total lengths of the RNA molecules. There are no hot spots of mutation for amino acid replacements of the influenza A strains, not even at the overlapping reading frames.

Comparing the NS1 proteins of PR8, A/FPV/Rostock/34 and A/Udorn/72 (Winter *et al.*, 1981), 24 amino acid replacements were found when FPV was compared with PR8. In 16 of these positions, the amino acids in the PR8 and Udorn

NS1 proteins are identical. On the other hand, 26 amino acid replacements were recognized, when the NS1 protein of the PR8 strain was compared with that of the Udorn strain. In 18 of these positions the amino acids in the FPV and PR8 NS1 proteins are identical. In the other 8 positions there are replacements between PR8 and Udorn, whereas the amino acids in corresponding positions in the NS1 protein of the Udorn strain are identical to those of the FPV strain. On the other hand, there are 35 amino acid replacements, when the NS1 proteins of FPV and Udorn are compared. These observations might indicate, that there is (1) some kind of a genetic drift within the NS1 proteins, similar to that found with the surface glycoproteins HA and NA; and that (2) the PR8 and FPV NS1 proteins might have a common ancestor, while the Udorn protein might be a drift product of (a close progenitor of) the PR8 gene (see Fig. 4), a hypothesis which is in agreement with the partial sequence studies by Hall and Air (1981). A similar but less convincing relationship is found for the NS2 proteins.

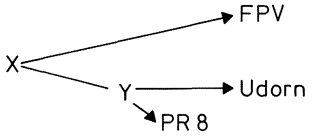


Fig. 4. Scheme of an evolutionary tree using amino acid changes in the NS1 proteins of the influenza strains A/FPV/Rostock/34, A/PR/8/34, and A/Udorn/72. X is regarded as a common ancestor of FPV and Y, while Y is regarded as a progenitor of the PR8 and Udorn strains

RNA segment 8 of PR8 has been sequenced in two different laboratories (Baez *et al.*, 1980; Winter *et al.*, 1981). In the deduced NS1 protein sequence a difference of one amino acid, and in the NS2 protein differences in two amino acids are seen, perhaps owing to different passage histories in the two laboratories.

A comparison of the nucleotide sequences and the deduced amino acid sequences of the NS genes of the PR8 (H1N1) and A/duck/Alberta/60/76 (H12N5) strains reveals significant sequence divergence (Baez *et al.*, 1981). The overall difference in nucleotide sequence was found to be 27.3%, while the differences were only 8 to 11% among the NS genes of the other 3 influenza A virus strains. When the NS1 and NS2 proteins of the PR8 and the A/duck/Alberta/60/76 strains were compared, it was found that the NS1 polypeptides are less conserved (33% difference) than the NS2 polypeptides (18.2% difference). This high degree of variation in the NS1 protein is observed over the total length of the molecule, including the region of the overlap of the two reading frames.

An exact comparison between the nucleotide or amino acid sequences of influenza A and B viruses for the NS1 and NS2 proteins is complicated by the different sizes of the genes. However, at a first approximation there is little or no sequence homology (Briedis and Lamb, 1982).

III. Genetic Relatedness of Viral Proteins

A. Differences in Migration Rates of Viral Proteins As Revealed by Polyacrylamide Gel Electrophoresis

When the proteins of various influenza virus strains are analyzed by polyacrylamide gel electrophoresis, marked differences in migration rates of corresponding proteins can be recognized. These variations are due to differences in molecular weight, charge or—for the glycoproteins—number and type of oligosaccharide side chains. The observed differences in migration rates of corresponding proteins were helpful in identifying the derivation of proteins in recombinant viruses. This information, in combination with the RNA analysis of recombinants was used to establish a genetic map of influenza viruses (for a review see Palese, 1977). A thorough analysis comparing the various proteins of avian influenza A virus subtypes has been published by Bosch *et al.* (1979).

Comparative protein studies of different influenza A and B viruses have been performed by Oxford *et al.* (1981) and Hugentobler *et al.* (1981). Of a total of 17 H1N1 isolates obtained between 1934 and 1953, 16 exhibited different migration patterns of the NP protein following electrophoresis in polyacrylamide gels. Seven isolates showed differences in the migration of the NS1 protein and the hemagglutinin (HA). Similar variations were observed when H3N2 strains isolated during the period 1968 to 1979 were compared. H1N1 strains isolated within shorter intervals (in the winters 1977/78 or 1978/79) showed less variation. But variations could be seen even in isolates obtained from single school and city outbreaks (Oxford *et al.*, 1981). Protein differences were also observed with 38 influenza B viruses isolated in epidemics from 1978 to 1980 (Hugentobler *et al.*, 1981).

Since these isolates were not further characterized by fingerprint or sequence analyses a definitive genetic comparison cannot be made. Presumably these differences in migration rates are due in each case to a few point mutations.

B. Tryptic Peptide Mapping

Trypsin cleaves a protein after arginine and lysine, leading to a mixture of peptides ending with these amino acids. The peptides can be separated by electrophoresis in one dimension and by chromatography in the second dimension. Using this method, two H1 strains (BEL and MEL) were compared in 1964 by Laver (1964). He was able to demonstrate that the hemagglutinins showed greater differences than the nucleoproteins. In another study, the hemagglutinins of human influenza virus strains isolated immediately before the appearance of the Hong Kong strain in 1968 were compared with HAs of strains isolated after 1968. The peptide maps of the HAs of the earlier isolates were similar, as were those of the later strains, but few similarities were observed between the peptide maps of the HAs of an early and a

late strain. This observation was taken as evidence that the Hong Kong virus was formed by reassortment (Laver and Webster, 1972). The hemagglutinins, especially the HA_2 subunits, of the duck Ukraine, the Equine 2 and the Hong Kong viruses showed a high degree of similarity in their peptide maps, indicating that their hemagglutinin genes had a common ancestor (Laver and Webster, 1973) (see also above).

Antigenic mutants were isolated by growing two H1 strains (swine and BEL) in the presence of low avidity antibodies. The peptide maps of such mutants revealed differences in the position of at least one peptide when compared to the wild-type strain (Laver and Webster, 1968). When monoclonal antibodies became available, antigenic variants were selected from the PR8 (H1N1) and the A/Mem/1/71 (H3N2) strains. These variants were expected to have only one or a few amino acid changes. Specific amino acid replacements were observed in the HA_1 subunits of 10 isolates, whereas no changes were found in their HA_2 subunits. Specifically, four independent variants selected with the same monoclonal antibody had the same amino acid replacement (Asp to Lys) (Laver *et al.*, 1978 a). Eight of 10 variants of the PR8 strain showed a single peptide change. In those cases which were analyzed, a serine in the wild-type HA was replaced by a leucine (Laver *et al.*, 1979 b). However, it has to be kept in mind that in this earlier study only one-third of the peptides, those which can be solubilized, were analyzed. Later on, when the total sequence of the hemagglutinin of an H3 strain was available, the sequence changes could be localized precisely. In ten variants of an H3N2 strain selected with monoclonal antibodies, the proline residues at position 143 in HA_1 changed to serine, threonine, leucine, or histidine, respectively. In other variants, an asparagine in position 133 changed to lysine, glycine 144 to aspartic acid, or serine 145 to lysine (Laver *et al.*, 1981). Residues 142 to 146 also changed in field strains of H3N2 influenza virus isolated between 1968 and 1977. However, other replacements were found as well (Laver *et al.*, 1980 a, 1980 b). These data suggest that one of the important antigenic sites of the HA molecule is located at positions 142 to 146. However, it cannot be excluded that amino acid replacements at these positions might alter the conformation of the polypeptide in such a way that new antigenic sites might become available which had been buried before. On the other hand, it has been shown that these amino acids are located at a rather exposed position (Wiley *et al.*, 1981; Wilson *et al.*, 1981).

The method of peptide mapping was also applied to a comparison of the NP and M proteins of 16 human and 3 lower vertebrate influenza A virus strains. By using shared peptides as taxonomic criteria, all nucleoproteins were grouped in one of two basic patterns, although some additional peptide differences existed within each NP group. There was no correlation between the NP groups and the hemagglutinin subtypes. This has been interpreted to mean that reassortment among different influenza virus strains has occurred (Dimmock *et al.*, 1980).

Five influenza B viruses isolated in 1940, 1970, 1973, 1976, and 1977 were analyzed by one-dimensional peptide mapping after digestion with *Staphylococcus aureus* V8 protease. The patterns of the membrane proteins were virtually identical, while the pattern of the nucleoprotein of the isolate from 1940 could be differentiated from those of the later strains. The most variable protein was the hem-

agglutinin. Again it appears that HA_1 is more variable than the HA_2. The neuraminidases also showed strain-specific differences in their patterns (Nakamura *et al.*, 1981).

C. Direct Sequencing

Direct amino acid sequencing is more tedious and material-consuming than the determination of the RNA sequence. However, without knowledge of the N-terminal and C-terminal sequences of a protein there is always uncertainty about whether the first start codon of an mRNA is indeed used to initiate the synthesis of that protein, or whether the first stop codon in an open reading frame indeed terminates the synthesis of that protein. Furthermore, amino acid sequencing aids in identifying any post-translational modifications in viral proteins.

So far, only the two glycoproteins HA and NA have been studied by direct amino acid sequencing. For the hemagglutinin it was found that—as mentioned above—its synthesis starts with a leader sequence, since the first 15 to 18 amino acids expected after the first start codon of the mRNA are not present at the N terminus of HA_1. Further properties of the HA molecule as revealed by direct sequencing have been discussed above (Skehel and Waterfield, 1975).

The cleavage sites of the hemagglutinins of three different subtypes (H1, H3, and H10) have been studied following cleavage *in vivo* or cleavage by trypsin or other proteases. After trypsin treatment of highly purified viruses, the HA_1 subunits lack a C-terminal arginine. It is speculated that the virus and not the host furnishes an exopeptidase of the carboxypeptidase B type that accomplishes this excision (Garten *et al.*, 1981). At the Ha_1/HA_2 cleavage site of the fowl plague virus hemagglutinin, an intervening sequence of 6 amino acids (5 basic ones) is removed (Garten *et al.*, 1982). This has been correlated with the pathogenic properties of the virus, and its ability to multiply in many different types of host cells (Bosch *et al.*, 1979).

The only influenza virus protein almost completely analyzed by direct amino acid sequencing is the hemagglutinin of the A/Memphis/102/72 (H3N2) strain (Dopheide and Ward, 1978 a, b, 1979, 1980 a, 1980 b, 1981; Ward and Dopheide, 1979, 1980). At the cleavage site between the carboxyl terminus of HA_1 and the amino terminus of HA_2, an arginine residue is removed during cleavage *in vivo*. Partial sequences of the hemagglutinins of other influenza virus strains have been determined by corresponding methods (Skehel and Waterfield, 1975; Bucher *et al.*, 1976; Waterfield *et al.*, 1977; McCauley *et al.*, 1977; Skehel *et al.*, 1980; Laver *et al.*, 1980 a, 1980 b; Ward *et al.*, 1981; Garten *et al.*, 1981, 1982). A comparison of extended amino acid sequences at the N termini of HA_1 and HA_2 of the Hong Kong (H3N2) and duck Ukraine strains (H3N2) has revealed their close relationship (Ward *et al.*, 1981), which is in accordance with the hybridization data (Scholtissek *et al.*, 1978) and sequencing of their RNAs (Fang *et al.*, 1981).

Partial amino acid sequence studies of four neuraminidase subtypes revealed that the protein lacks a signal peptide and is anchored in the viral membrane by the

hydrophobic domain at the N terminus. It appears that the N terminus of the neuraminidase is not modified following its synthesis (Blok *et al.*, 1982). In this respect, the neuraminidase differs from other viral spike proteins, such as the hemagglutinin (see above) or the VSV G protein (Rose and Gallione, 1981).

IV. Concluding Remarks

Since influenza viruses have a segmented genome, each RNA segment can function independently, as shown by stepwise chemical inactivation (Scholtissek and Rott, 1964). Therefore, new influenza virus strains can be created by reassortment either under natural conditions or in tissue culture. This implies that, when two different influenza viruses are compared genetically, some RNA segments (or proteins, respectively) might be—except for a new mutations—identical, while others exhibit relatively little homology. This is exactly what has been found. Reassortment, especially for the genes coding for the surface glycoproteins hemagglutinin and neuraminidase, seems to be the mechanism for the creation of new pandemic strains (antigenic shift).

The RNA segments coding for the internal proteins (P-proteins, NP, M) are relatively conserved, while the RNA segments coding for the glycoproteins (HA, NA) exhibit a relatively high variation. Some of the variable domains of the hemagglutinin genes can be correlated with the antigenic sites of the gene products. The mutation rate seems to be similar for the various RNA segments and even within the segments, as determined by the scatter of silent mutations along the RNA molecules. Therefore, the increased amino acid replacements forming the variable regions seem to be due to mutation and selection by the immune system of the host. This, however, implies that these variable domains are not too important for the function of the gene products; *i.e.* the cell attachment site or enzymatic activity, respectively.

There is another kind of variation found with the genes coding for the glycoproteins: While the other proteins have almost identical molecular weights, when different subtypes are compared, the glycoproteins of different subtypes can vary in size. Especially for the neuraminidase, extended deletions and (or) insertions have been recognized.

Thirteen different subtypes of the hemagglutinin and 9 subtypes of the neuraminidase of influenza A viruses can be discerned. Within each subtype there is certain variation. However, this variation does not exceed that between the various subtypes, although some subtypes are more closely related to each other than others (Webster *et al.*, 1982; see Fig. 3). A comparison of the conserved domains of the hemagglutinins and neuraminidases suggests that the various subtypes are all derived from a common ancestor molecule. It is not yet understood why only a limited number of subtypes have evolved.

In terms of the genetic relatedness of the nonstructural proteins, the influenza A viruses can be divided into two groups. The derivation of these two groups, and their underlying significance is not yet understood. A genetic "drift" has been recognized also with these genes; however, the selection principle—if there is any—is not yet known.

Acknowledgements

I am very much indebted to Drs. R. R. Friis and R. Rott for excellent help and discussion during the preparation of the manuscript, and especially to Mrs. M. Seitz for her patience during typing of the manuscript.

Experiments carried out in the author's laboratory were supported by the Sonderforschungsbereich 47 of the Deutsche Forschungsgemeinschaft, and by the Fonds der Chemischen Industrie.

References

Aaronson, P. P., Young, J. F., Palese, P.: Oligonucleotode mapping: evaluation of its sensitivity by computer-simulation. Nucl. Acids Res. *10,* 237–246 (1982).

Air, G.: Nucleotide sequence coding for the "signal peptide" and N terminus of the hemagglutinin of an Asian (H2N2) strain of influenza virus. Virology *97,* 468–472 (1979).

Air, G. M.: Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus. Proc. Natl. Acad. Sci. U.S.A. *78,* 7639–7643 (1981).

Air, G. M., Hackett, J. A.: Gene 8 of influenza virus: sequences of cDNA transcribed from the 3′ ends of viral RNA of influenza A and B strains. Virology *103,* 291–298 (1980).

Allen, H., McCauley, J., Waterfield, M., Gething, M.-J.: Influenza virus RNA segment 7 has the coding capacity for two polypeptides. Virology *107,* 548–551 (1980).

Almond, J. W., Barry, R. D.: Genetic recombination between two strains of fowl plague virus: construction of genetic maps. Virology *92,* 407–415 (1979).

Baez, M., Taussig, R., Zazra, J. J., Young, J. F., Palese, P.: Complete nucleotide sequence of the influenza A/PR/8/34 virus NS gene and comparison with the NS genes of the A/Udorn/72 and A/FPV/Rostock/34 strains. Nucl. Acids Res. *8,* 5845–5858 (1980).

Baez, M., Zazra, J. J., Elliott, R. M., Young, J. F., Palese, P.: Nucleotide sequence of the influenza A/duck/Alberta/60/76 virus NS RNA: conservation of the NS1/NS2 overlapping gene structure in a divergent influenza virus RNA segment. Virology *113,* 397–402 (1981).

Bean, W. J., Sriram, G., Webster, R. G.: Electrophoretic analysis of iodine-labeled influenza virus RNA segments. Analyt. Biochem. *102,* 228–232 (1980 a).

Bean, W. J., Cox, N. J., Kendal, A. P.: Recombination of human influenza A viruses in nature. Nature *284,* 638–640 (1980 b).

Bean, W. J., Hinshaw, V. S., Webster, R. G.: Genetic characterization of an influenza virus from seals. In: The Replication of Negative Strand Viruses (Bishop, D. H. L., Compans, R. W., eds.), 363–367. New York-Amsterdam-Oxford: Elsevier/North-Holland 1981.

Bishop, D. H. L., Huddleston, J. A., Brownlee, G. G.: The complete sequence of RNA segment 2 of influenza A/NT/60/68 and its encoded P1 protein. Nucl. Acids Res. *10,* 1335–1343 (1982).

Blok, J., Air, G. M.: Comparative nucleotide sequences at the 3′ end of the neuraminidase gene from eleven influenza A viruses. Virology *107,* 50–60 (1980).

Blok, J., Air, G. M.: Block deletions in the neuraminidase genes from some influenza A viruses of the N1 subtype. Virology *118,* 229–234 (1982).

Blok, J., Air, G. M., Laver, W. G., Ward, C. W., Lilley, G. G., Woods, E. F., Roxburgh, C. M., Inglis, A. S.: Studies on the size, chemical composition, and partial sequence of the neuraminidase (NA) from type A influenza viruses show that the N terminal region of the NA is not processed and serves to anchor the NA in the viral membrane. Virology *119,* 109–121 (1982).

Bosch, F. X., Orlich, M., Klenk, H.-D., Rott, R.: The structure of the hemagglutinin, a determinant for the pathogenicity of influenza viruses. Virology *95,* 197–207 (1979).

Bosch, F. X., von Hoyningen-Huene, V., Scholtissek, C., Rott, R.: The overall evolution of the H7 influenza virus hemagglutinin is different from the evolution of the proteolytic cleavage site. J. Gen. Virol. *61,* 101–104 (1982).

Both, G. W., Sleigh, M. J.: Complete nucleotide sequence of the hemagglutinin gene from a human influenza virus of the Hong Kong subtype. Nucl. Acids Res. *8,* 2561–2575 (1980).

Both, G. W., Sleigh, M. J.: Conservation and variation in the hemagglutinins of Hong Kong subtype influenza viruses during antigenic drift. J. Virol. *39,* 663–672 (1981).

Both, G. W., Sleigh, M. J., Bender, V. J., Moss, B. A.: A comparison of antigenic variation in Hong Kong influenza virus hemagglutinins at the nucleic acid level. In: Structure and Variation in Influenza Virus (Laver, G., Air, G., eds.), 81–89. New York-Amsterdam-Oxford: Elsevier/North-Holland 1980.

Briedis, D. J., Lamb, R. A.: Influenza B virus genome: Sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS1 and NS2 proteins. J. Virol. *42,* 186–193 (1982).

Briedis, D. J., Lamb, R. A., Choppin, P. W.: Sequence of RNA segment 7 of the influenza B virus genome: partial amino acid homology between the membrane proteins (M1) of influenza A and B viruses and conservation of a second open reading frame. Virology *116,* 581–588 (1982).

Bucher, D. J., Li, S. S.-L., Kehoe, J. M., Kilbourne, E. D.: Chromatographic isolation of the hemagglutinin polypeptides from influenza virus vaccine and determination of their amino-terminal sequences. Proc. Natl. Acad. Sci. U.S.A. *78,* 238–242 (1976).

Desselberger, U., Palese, P.: Molecular weights of RNA segments of influenza A and B viruses. Virology *88,* 394–399 (1978).

Desselberger, U., Nakajima, K., Alfino, P., Pedersen, F. S., Haseltine, W. A., Hannoun, C., Palese, P.: Biochemical evidence that "new" influenza virus strains in nature may arise by recombination (reassortment). Proc. Natl. Acad. Sci. U.S.A. *75,* 3341–3345 (1978).

Desselberger, U., Racaniello, V. R., Zazra, J. J., Palese, P.: The 3′ and 5′-terminal sequences of influenza A, B, and C virus RNA segments are highly conserved and show partial inverted complementarity. Gene *8,* 315–328 (1980).

De Wachter, R., Fiers, W.: Preparative two-dimensional polyacrylamide gel electrophoresis of ^{32}P-labeled RNA. Analyt. Biochem. *49,* 184–197 (1972).

Dimmock, N. J., Carver, A. S., Webster, R. G.: Categorization of nucleoproteins and matrix proteins from type A influenza viruses by peptide mapping. Virology *103,* 350–356 (1980).

Donis-Keller, H., Maxam, A. M., Gilbert, W.: Mapping adenines, guanines, and pyrimidines in RNA. Nucl. Acids Res. *4,* 2527–2538 (1977).

Dopheide, T. A. A., Ward, C. W.: Studies on the primary structure of the hemagglutinin from an H3N2 variant A/Memphis/102/72. In: Topics in Infectious Diseases, Vol. 3: The Influenza Virus Hemagglutinin (Laver, W. G., Bachmayer, H., Weil, R., eds.), 193–201. Wien-New York: Springer 1978 a.

Dopheide, T. A., Ward, C. W.: The carboxyl-terminal sequence of the heavy chain of a Hong Kong influenza hemagglutinin. Eur. J. Biochem. *85,* 393–398 (1978 b).

Dopheide, T. A. A., Ward, C. W.: A Hong Kong influenza hemagglutinin light chain: The amino acid sequence of cyanogen bromide fragments CN3 and CN4 and the N terminal 45 residues. Virology *92,* 230–235 (1979).

Dopheide, T. A., Ward, C. W.: Structural studies on a Hong Kong hemagglutinin. The structure of the light chain and the arrangement of the disulphide bonds. In: Structure and Variation in Influenza Virus (Laver, G., Air, G., eds.), 21–26. New York-Amsterdam-Oxford: Elsevier/North-Holland 1980 a.

Dopheide, T. A., Ward, C. W.: The amino acid sequence of a Hong Kong influenza hemagglutinin light (HA2) chain. J. Gen. Virol. *50,* 329–335 (1980 b).

Dopheide, T. A., Ward, C. W.: The location of the bromelain cleavage site in a Hong Kong influenza hemagglutinin. J. Gen. Virol. *52,* 362–370 (1981).

Emtage, J. S., Catlin, G. H., Carey, N. H.: Polyadenylation and reverse transcription of influenza viral RNA. Nucl. Acids Res. *6,* 1221–1239 (1979).

Fang, R., Min Jou, W., Huylebroeck, D., Devos, R., Fiers, W.: Complete structure of A/duck/Ukraine/63 influenza hemagglutinin gene: Animal virus as progenitor of human H3 Hong Kong 1968 influenza hemagglutinin. Cell *25,* 315–323 (1981).

Fields, S., Winter, G.: Nucleotide sequence of influenza virus segments 1 and 3 reveal mosaic structure of a small viral RNA segment. Cell *28,* 303–313 (1982).

Fields, S., Winter, G., Brownlee, G. G.: Structure of the neuraminidase gene in human influenza virus A/PR/8/34. Nature *290,* 213–217 (1981).

Frisby, D.: Oligonucleotide mapping of non-radioactive virus and messenger RNA. Nucl. Acids Res. *4,* 2975–2996 (1977).

Garten, W., Bosch, F.X., Linder, D., Rott, R., Klenk, H.-D.: Proteolytic activation of the influenza virus hemagglutinin: The structure of the cleavage site and the enzymes involved in cleavage. Virology *115,* 361–374 (1981).

Garten, W., Linder, D., Rott, R., Klenk, H.-D.: The cleavage site of the hemagglutinin of fowl plague virus. Virology *121,* 186–190 (1982).

Gething, M.-J., Bye, J., Skehel, J.J., Waterfield, M.: Cloning and DNA sequence of double-stranded copies of haemagglutinin genes from H2 and H3 strains elucidates antigenic shift and drift in human influenza virus. Nature *287,* 301–306 (1980).

Ghendon, Y., Klimov, A., Blagoveshenskaya, O., Genkina, D.: Investigation of recombinants of human influenza and fowl plague viruses. J. Gen. Virol. *43,* 183–191 (1979).

Ghendon, Y.Z., Klimov, A.I., Alexandrova, G.I., Polezhaev, F.I.: Analysis of genome composition and reactogenicity of recombinants of cold-adapted and virulent virus strains. J. Gen. Virol. *53,* 215–224 (1981 a).

Ghendon, Y., Klimov, A., Gorodkova, N., Döhner, L.: Genome analysis of influenza A virus strains isolated during an epidemic of 1979–1980. J. Gen. Virol. *56,* 303–313 (1981 b).

Hall, R.M., Air, G.M.: Variation in nucleotide sequences coding for the N terminal regions of the matrix and nonstructural proteins of influenza A viruses. J. Virol. *38,* 1–7 (1981).

Harms, E., Rohde, W., Bosch, F., Scholtissek, C.: Biochemical studies on influenza viruses. II. Assignment of gene functions to RNA segments 5, 7, and 8 of fowl plague virus and virus N. Virology *86,* 413–422 (1978).

Hay, A.J., Bellamy, A.R., Abraham, G., Skehel, J.J., Brand, C.M., Webster, R.G.: Procedure for characterization of genetic material of candidate vaccine strains. (Int. Symp. Influenza Immunization, 2nd, 1977.) Dev. Biol. Stand. *39,* 15–24 (1977).

Heller, E., Scholtissek, C.: Evidence for intracistronic complementation of the product of the influenza virus gene Ptra (P3 of fowl plague virus). J. Gen. Virol. *49,* 133–139 (1980).

Hinshaw, V.S., Bean, W.J., Webster, R.G., Easterday, B.C.: The prevalence of influenza viruses in swine and the antigenic and genetic relatedness of influenza viruses from man and swine. Virology *84,* 51–62 (1978).

Hinshaw, V.S., Bean, W.J., Webster, R.G., Sriram, G.: Genetic reassortment of influenza A viruses in the intestinal tract of ducks. Virology *102,* 412–419 (1980).

Hinshaw, V.S., Air, G.M., Gibbs, A.J., Graves, L., Prescott, B., Karunakaran, D.: Antigenic and genetic characterization of a novel hemagglutinin subtype of influenza A viruses from gulls. J. Virol. *42,* 865–872 (1982).

Hiti, A.L., Nayak, D.P.: Complete nucleotide sequence of the neuraminidase gene of human influenza virus A/WSN/33. J. Virol. *41,* 730–734 (1982).

Hiti, A.L., Davis, A.R., Nayak, D.P.: Complete sequence analysis shows that the hemagglutinins of the H0 and H2 subtypes of human influenza virus are closely related. Virology *111,* 113–124 (1981).

Horisberger, M.A.: The large P proteins of influenza A viruses are composed of one acidic and two basic polypeptides. Virology *107,* 302–305 (1980).

Horisberger, M.A.: Identification of a catalytic activity of the large basic P polypeptide of influenza virus. Virology *120,* 279–286 (1982).

Huddleston, J.A., Brownlee, G.G.: The sequence of the nucleoprotein gene of human influenza A virus strain A/NT/60/68. Nucl. Acids Res. *10,* 1029–1038 (1982).

Hugentobler, A.L., Schild, G.C., Oxford, J.S.: Differences in the electrophoretic migration rates of polypeptides and RNAs of recent isolates of influenza B viruses. Arch. Virol. *69,* 197–207 (1981).

Kaptein, J.S., Nayak, D.P.: Complete nucleotide sequence of the polymerase 3 gene of human influenza virus A/WSN/33. J. Virol. *42,* 55–63 (1982).

Klimov, A.I., Ghendon, Y.Z.: Genome analysis of H1N1 influenza virus strains isolated in the USSR during an epidemic in 1961–1962. Arch. Virol. *70,* 225–232 (1981).

Kozlov, J.V., Gorbulev, V.G., Kurmanova, A.G., Bayev, A.A., Shilov, A.A., Zhdanov, V.M.: On the origin of the H1N1 (A/USSR/90/77) influenza virus. J. Gen. Virol. *56,* 437–440 (1981).

Krug, R.M.: The unique interaction of influenza viral RNA transcription with the host-cell transcriptional machinery. In: The Origin of Pandemic Influenza Viruses (Laver, W.G., ed.). New York-Amsterdam-Oxford: Elsevier/North-Holland 1983 (in press).

Lamb, R. A., Lai, C.-J.: Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. Cell *21,* 475–485 (1980).

Lamb, R. A., Choppin, P. W.: Identification of a second protein (M2) encoded by RNA segment 7 of influenza virus. Virology *112,* 729–737 (1981).

Lamb, R. A., Lai, C.-J.: Conservation of the influenza virus membrane protein (M1) amino acid sequence and an open reading frame of RNA segment 7 encoding a second protein (M2) in H1N1 and H3N2 strains. Virology *112,* 746–751 (1981).

Laver, W. G.: Structural studies on the protein subunits from three strains of influenza virus. J. Mol. Biol. *9,* 109–124 (1964).

Laver, W. G., Webster, R. G.: Selection of antigenic mutants of influenza viruses. Isolation and peptide mapping of their hemagglutinating proteins. Virology *34,* 193–202 (1968).

Laver, W. G., Webster, R. G.: Studies on the origin of pandemic influenza. II. Peptide maps of the light and heavy polypeptide chains from the hemagglutinin subunits of A2 influenza viruses isolated before and after the appearance of Hong Kong influenza. Virology *48,* 445–455 (1972).

Laver, W. G., Webster, R. G.: Studies on the origin of pandemic influenza. III. Evidence implicating chick and equine influenza viruses as possible progenitors of the Hong Kong strain of human influenza. Virology *51,* 383–391 (1973).

Laver, W. G., Air, G. M., Webster, R. G., Gerhard, W., Ward, C. W., Dopheide, T. A. A.: Antigenic drift in type A influenza virus: Sequence differences in the hemagglutinin of Hong Kong (H3N2) variants selected with monoclonal hybridoma antibodies. Virology *98,* 226–237 (1979 a).

Laver, W. G., Gerhard, W., Webster, R. G., Frankel, M. E., Air, G. M.: Antigenic drift in type A influenza virus: peptide mapping and antigenic analysis of A/PR/8/34 (H0N1) variants selected with monoclonal antibodies. Proc. Natl. Acad. Sci. U.S.A. *76,* 1425–1429 (1979 b).

Laver, W. G., Air, G. M., Dopheide, T. A., Ward, C. W.: Amino acid sequence changes in the hemagglutinin of A/Hong Kong (H3N2) influenza virus during the period 1968–1977. Nature *283,* 454–457 (1980 a).

Laver, W. G., Air, G. M., Webster, R. G., Gerhard, W., Ward, C. W., Dopheide, T. A. A.: The mechanism of antigenic drift in influenza virus: sequence changes in the hemagglutinin of variants selected with monoclonal hybridoma antibodies. Phil. Trans. R. Soc. (Lond.) *B 288,* 313–326 (1980 b).

Laver, W. G., Air, G. M., Webster, R. G.: Mechanism of antigenic drift in influenza virus. Amino acid sequence changes in an antigenically active region of Hong Kong (H3N2) influenza virus hemagglutinin. J. Mol. Biol. *145,* 339–361 (1981).

Markoff, L., Lai, C.-J.: Sequence of the influenza A/Udorn/72 (H3N2) virus neuraminidase gene as determined from cloned full-length DNA. Virology *119,* 288–297 (1982).

Maxam, A. M., Gilbert, W.: A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. *74,* 560–564 (1977).

McCauley, J., Skehel, J. J., Waterfield, M. D.: The structure of the small polypeptide chain of the hemagglutinin of an Asian influenza virus Japan 305/57-Bellamy/42 (H2N1). In: Topics in Infectious Diseases, Vol. 3: The Influenza Virus Hemagglutinin (Laver, W. G., Bachmayer, H., Weil, R., eds.), 181–192. Wien-New York: Springer 1978.

McCauley, J. W., Mahy, B. W. J., Inglis, S. C.: Nucleotide sequence of fowl plague virus RNA segment 7. J. Gen. Virol. *58,* 211–215 (1982).

McGeoch, D., Fellner, P., Newton, C.: Influenza virus genome consists of eight distinct RNA species. Proc. Natl. Acad. Sci. U.S.A. *73,* 3045–3049 (1976).

Meier-Ewert, H., Petri, T., Bishop, D. H. L.: Oligonucleotide fingerprint analyses of influenza C virion RNA recovered from five different isolates. Arch. Virol. *67,* 141–147 (1981).

Min Jou, W., Verhoeyen, M., Devos, R., Saman, E., Fang, R., Huylebroeck, D., Fiers, W.: Complete structure of the hemagglutinin gene from the human influenza A/Victoria/3/75 (H3N2) strain as determined from cloned DNA. Cell *19,* 683–696 (1980).

Nakajima, K., Desselberger, U., Palese, P.: Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. Nature *274,* 334–339 (1978).

Nakajima, K., Nakajima, S., Sugiura, A.: The possible origin of H3N2 influenza virus. Virology *120,* 504–509 (1982).

Nakamura, K., Kitame, F., Homma, M.: A comparison of proteins among various influenza B virus strains by one-dimensional peptide mapping. J. Gen. Virol. *56,* 315–323 (1981).

Nerome, K., Ishida, M., Kakayama, M., Oya, A., Kanai, C., Suwicha, K.: Antigenic and genetic analysis of A/Hong Kong (H3N2) influenza viruses isolated from swine and man. J. Gen. Virol. *56*, 441–445 (1981).

Ortin, J., Najera, R., Lopez, C., Davila, M., Domingo, E.: Genetic variability of Hong Kong (H3N2) influenza viruses: spontaneous mutations and their location in the viral genome. Gene *11*, 319–331 (1980).

Oxford, J. S., Corcoran, T., Schild, G. C.: Intratypic electrophoretic variation of structural and non-structural polypeptides of human influenza A viruses. J. Gen. Virol. *56*, 431–436 (1981).

Palese, P.: The genes of influenza virus. Cell *10*, 1–10 (1977).

Palese, P., Schulman, J. L.: Mapping of the influenza virus genome: Identification of the hemagglutinin and the neuraminidase genes. Proc. Natl. Acad. Sci. U.S.A. *73*, 2142–2146 (1976).

Palese, P., Ritchey, M. B., Schulman, J. L.: Mapping of the influenza virus genome. II. Identification of the P1, P2, and P3 genes. Virology *76*, 114–121 (1977 a).

Palese, P., Ritchey, M. B.: Polyacrylamide gel electrophoresis of the RNA of new influenza virus strains: An epidemiological tool. (Int. Symp. Influenza Immunization, 2nd, 1977.) Dev. Biol. Stand. *39*, 411–415 (1977 b).

Palese, P., Racaniello, V. R., Desselberger, U., Young, J., Baez, M.: Genetic structure and genetic variation of influenza viruses. Phil. Trans. R. Soc. (Lond.) *B 288*, 299–305 (1980).

Palese, P., Elliott, R. M., Baez, M., Zazra, J. J., Young, J. F.: Genome diversity among influenza A, B, and C viruses and genetic structure of RNA 7 and RNA 8 of influenza A viruses. In: Genetic Variation among Influenza Viruses (Nayak, D. P., ed.) (ICN-UCLA Symposia on Molecular and Cellular Biology. Vol. XXI), 127–140. New York-London-Toronto-Sydney-San Francisco: Academic Press 1981.

Porter, A. G., Barber, C., Carey, N. H., Hallewell, R. A., Threlfall, G., Emtage, J. S.: Complete nucleotide sequence of an influenza virus hemagglutinin gene from cloned DNA. Nature *282*, 471–477 (1979).

Porter, A. G., Smith, J. C., Emtage, J. S.: Nucleotide sequence of influenza virus RNA segment 8 indicates that coding regions for NS1 and NS2 proteins overlap. Proc. Natl. Acad. Sci. U.S.A. *77*, 5074–5078 (1980).

Ritchey, M. B., Palese, P., Schulman, J. L.: Mapping of the influenza virus genome. III. Identification of genes coding for nucleoproteins, membrane protein, and nonstructural protein. J. Virol. *20*, 307–313 (1976).

Robertson, J. S.: 5′ and 3′ terminal nucleotide sequences of the RNA genome segments of influenza virus. Nucl. Acids Res. *6*, 3745–3757 (1979).

Rohde, W., Harms, E., Scholtissek, C.: Biochemical studies on influenza viruses. I. Comparative analysis of equine 2 virus and virus N genes and gene products. Virology *79*, 393–404 (1977).

Rose, J. K., Gallione, C. J.: Nucleotide sequences of the mRNA's encoding the vesicular stomatitis virus G and M proteins determined from cDNA clones containing the complete coding regions. J. Virol. *39*, 519–528 (1981).

Sanger, F., Nicklen, S., Coulson, A. R.: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. *74*, 5463–5467 (1977).

Scholtissek, C.: Synthesis in vitro of RNA complementary to parental viral RNA by RNA polymerase induced by influenza virus. Biochim. Biophys. Acta *179*, 389–397 (1969).

Scholtissek, C.: The genome of the influenza virus. Curr. Top. Microbiol. Immunol. *80*, 139–169 (1978).

Scholtissek, C.: Influenza virus genetics. Adv. Genetics *20*, 1–36 (1979 a).

Scholtissek, C.: The genes coding for the surface glycoproteins of influenza A viruses contain a small conserved and a large variable region. Virology *93*, 594–597 (1979 b).

Scholtissek, C.: The structure of the hemagglutinin and neuraminidase genes as revealed by molecular hybridization. In: Structure and Variation in Influenza Virus (Laver, G., Air, G., eds.), 191–199. New York-Amsterdam-Oxford: Elsevier/North-Holland 1980.

Scholtissek, C., Rott, R.: Behavior of virus-specific activities in tissue cultures infected with myxoviruses after chemical changes of the viral ribonucleic acid. Virology *22*, 169–176 (1964).

Scholtissek, C., Rott, R.: Hybridization studies with influenza virus RNA. Virology *39*, 400–407 (1969).

Scholtissek, C., Rott, R.: Synthesis in vivo of influenza virus plus and minus strand RNA and its preferential inhibition by antibiotics. Virology *40,* 989–996 (1970).

Scholtissek, C., von Hoyningen-Huene, V.: Genetic relatedness of the gene which codes for the non-structural (NS) protein of different influenza A strains. Virology *102,* 13–20 (1980).

Scholtissek, C., Harms, E., Rohde, W., Orlich, M., Rott, R.: Correlation between RNA fragments of fowl plague virus and their corresponding gene functions. Virology *74,* 332–344 (1976).

Scholtissek, C., Rohde, W., Harms, E., Rott, R.: Correlation between the base sequence homology of RNA segment 4 and antigenicity of the hemagglutinin of influenza viruses. Virology *79,* 330–336 (1977 a).

Scholtissek, C., Rohde, W., Harms, E.: Genetic relationship between an influenza A and B virus. J. Gen. Virol. *37,* 243–247 (1977 b).

Scholtissek, C., Rohde, W., von Hoyningen, V., Rott, R.: On the origin of the human influenza virus subtypes H2N2 and H3N2. Virology *87,* 13–20 (1978 a).

Scholtissek, C., von Hoyningen, V., Rott, R.: Genetic relatedness between the new 1977 epidemic strains (H1N1) of influenza and human influenza strains isolated between 1947 and 1957 (H1N1). Virology *89,* 613–617 (1978 b).

Scholtissek, C., Rott, R., Lvov, D. K., Myasnikova, I. A.: Genetic relatedness of the neuraminidase of influenza A strains Nav2, Nav3, and Nav6. Arch. Virol. *65,* 325–328 (1980).

Skehel, J. J., Waterfield, M. D.: Studies on the primary structure of influenza virus hemagglutinin: Proc. Natl. Acad. Sci. U.S.A. *72,* 93–97 (1975).

Skehel, J. J., Hay, A. J.: Nucleotide sequences at the 5′ termini of influenza virus RNAs and their transcripts. Nucl. Acids Res. *5,* 1207–1219 (1978).

Skehel, J. J., Waterfield, M. D., McCauley, J. W., Elder, K., Wiley, D. C.: Studies on the structure of the hemagglutinin. Phil. Trans. R. Soc. (Lond.) *B 288,* 335–339 (1980).

Sleigh, M. J., Both, G. W., Underwood, P. A., Bender, V. J.: Antigenic drift in the hemagglutinin of the Hong Kong influenza subtype: correlation of amino acid changes with alterations in viral antigenicity. J. Virol. *37,* 845–853 (1981).

Sriram, G., Bean, W. J., Hinshaw, V. S., Webster, R. G.: Genetic diversity among avian influenza viruses. Virology *105,* 592–599 (1980).

Van Rompuy, L., Min Jou, W., Huylebroeck, D., Devos, R., Fiers, W.: Complete nucleotide sequence of the nucleoprotein gene from the human influenza strain A/PR/8/34 (H0N1). Eur. J. Biochem. *116,* 347–353 (1981). Corrections: Eur. J. Biochem. *126,* 645 (1982).

Verhoeyen, M., Fang, R., Min Jou, W., Devos, R., Huylebroeck, D., Saman, E., Fiers, W.: Antigenic drift between the hemagglutinin of the Hong Kong influenza strains A/Aichi/2/68 and A/Victoria/3/75. Nature *286,* 771–776 (1980).

Ward, C. W., Dopheide, T. A. A.: A Hong Kong influenza hemagglutinin light chain: Amino acid sequence of cyanogen bromide fragment CN2. Virology *95,* 107–118 (1979).

Ward, C. W., Dopheide, T. A.: The Hong Kong (H3) hemagglutinin. Complete amino acid sequence and oligosaccharide distribution for the heavy chain of A/Memphis/102/72. In: Structure and Variation in Influenza Virus (Laver, G., Air, G., eds.), 27–38. New York-Amsterdam-Oxford: Elsevier/North-Holland 1980.

Ward, C. W., Webster, R. G., Inglis, A. S., Dopheide, T. A.: Composition and sequence studies show that A/duck/Ukraine/1/63 hemagglutinin (Hav7) belongs to the Hong Kong (H3) subtype. J. Gen. Virol. *53,* 163–168 (1981).

Waterfield, M. D., Skehel, J. J., Nakashima, Y., Gurnett, A., Bilham, T.: Structural studies of the hemagglutinin of the Asian influenza virus Japan/305/57-Bellamy/42 (H2N1). Cyanogen bromide cleavage of the larger polypeptide chain HA1. In: Topics in Infectious Diseases, Vol. 3: The Influenza Virus Hemagglutinin (Laver, W. G., Bachmayer, H., Weil, R., eds.), 167–179. Wien-New York: Springer 1978.

Webster, R. G., Hinshaw, V. S., Bean, W. J., Van Wyke, K. L., Geraci, J. R., St. Aubin, D. J., Petursson, G.: Characterization of an influenza A virus from seals. Virology *113,* 712–724 (1981).

Webster, R. G., Laver, W. G., Air, G. M., Schild, G. C.: Molecular mechanisms of variation in influenza viruses. Nature *296,* 115–121 (1982).

WHO-Bulletin (1979). Reconsideration of influenza A virus nomenclature: a WHO Memorandum. Bull. World Health Org. *57,* 227–233 (1979).

WHO-Bulletin (1980). A revision of the system of nomenclature for influenza viruses: a WHO Memorandum. Bull. World Health Org. *58,* 585–591 (1980).

Wiley, D. C., Wilson, I. A., Skehel, J. J.: Structural identification of the antibody-binding site of Hong Kong influenza hemagglutinin and their involvement in antigenic variation. Nature *289,* 373–378 (1981).

Wilson, I. A., Skehel, J. J., Wiley, D. C.: Structure of the hemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature *289,* 366–373 (1981).

Winter, G., Fields, S.: Cloning of influenza cDNA into M13: the sequence of the RNA segment encoding the A/PR/8/34 matrix protein. Nucl. Acids Res. *8,* 1965–1974 (1980).

Winter, G., Fields, S.: The structure of the gene encoding the nucleoprotein of human influenza virus A/PR/8/34. Virology *114,* 423–428 (1981).

Winter, G., Fields, S.: Nucleotide sequence of human influenza A/PR/8/34 segment 2. Nucl. Acids Res. *10,* 2135–2143 (1982).

Winter, G., Fields, S., Gait, M. J., Brownlee, G. G.: The use of synthetic oligodeoxynucleotide primers in cloning and sequencing segment 8 of influenza virus (A/PR/8/34). Nucl. Acids Res. *9,* 237–245 (1981 a).

Winter, G., Fields, S., Brownlee, G. G.: Nucleotide sequence of the hemagglutinin gene of a human influenza virus H1 subtype. Nature *292,* 72–75 (1981 b).

Young, J. F., Palese, P.: Evolution of human influenza A viruses in nature: Recombination contributes to genetic variation of H1N1 strains. Proc. Natl. Acad. Sci. U.S.A. *76,* 6547–6551 (1979).

Young, J. F., Desselberger, U., Palese, P.: Evolution of human influenza A viruses in nature: Sequential mutations in the genomes of new H1N1 isolates. Cell *18,* 73–83 (1979).

5

Antigenic Variation Among Type A Influenza Viruses

R. G. Webster, W. G. Laver, and *G. M. Air*

I. Introduction

Influenza is a worldwide disease that periodically causes epidemics in humans, horses, pigs, birds, and occasionally in other animals such as seals. The hallmark of influenza virus is its variability, having the capacity to change its antigenic identity so remarkably that the specific immunity established in response to infection by a particular strain may give little or no protection against viruses that subsequently arise. Antigenic variation is the result of molecular changes in the surface proteins of influenza viruses, *i.e.,* the hemagglutinin (HA) and neuraminidase (NA), which reflect alterations in the sequences of the constituent genes. There are two types of change that the HA and NA can undergo; these occur by different mechanisms and are known as antigenic "drift" and antigenic "shift". Antigenic "drift" occurs within a subtype and involves a series of minor changes at the gene level (usually point mutations) so that variants are formed, each slightly different from its predecessor. On the other hand, antigenic "shift" is caused by a more radical change in the HA and/or NA. Here influenza viruses appear in the population with surface antigens unlike those of viruses immediately preceding them. The origin of these new viruses is still unclear; some may result from genetic reassortment between human and animal virus strains and others may exist in a dormant stage for long periods before reemerging.

Because of antigenic variation, it has proved impossible to control influenza by vaccination, and variations in virulence, host range, and transmissibility of the virus influence the spread and severity of influenza epidemics. This review summarizes recent information on the ways in which influenza genes and gene product vary and the effects of these changes on the biological activities of the virus.

II. Historical

The highly contagious, acute respiratory illness known as influenza appears to have afflicted humans since ancient times. The sudden appearance of epidemics of respiratory diseases that persist for a few weeks and equally suddenly disappear are sufficiently characteristic to permit identification of a number of major epidemics in the distant past. One such epidemic was recorded by Hippocrates in 412 B.C., and numerous episodes were described in the Middle Ages. Hirsch (1883) and Webster (1800) collected historical data from 1500 until the late nineteenth century which was reviewed by Noble (1982); some features emerge:

(1) Epidemics occurred relatively frequently but at irregular intervals and occasionally the disease seemed to disappear for periods of time.

(2) Epidemics varied in severity but usually caused mortality in the elderly.

(3) Some epidemics such as the ones in 1781 and 1830 appeared to spread across Russia from Asia.

There are some similarities with our current knowledge of the epidemiology of influenza; recent epidemics have appeared first in China and usually have had the greatest effect on the elderly. There is as yet no indication that influenza disappears from the human population but in recent years it has become relatively mild, causing little excess mortality.

Retrospective seroepidemiology suggests that the influenza epidemic in humans in 1889—90 was caused by a virus antigenically similar to the "Asian" strains (H2N?) (Mulder and Masurel, 1958). The epidemic of 1900 may have been caused by a "Hong Kong"-like strain, but with a neuraminidase like that of A/Equine/Miami/1/63 influenza virus. Swine-like influenza virus (H1N1) is believed to have caused the catastropic pandemic in 1918—19 that resulted in excess of 20 million deaths worldwide (Davenport *et al.*, 1953, 1964; Masurel, 1968). The seroepidemiological information suggested that human strains recirculate and the reappearance of the H1N1 strain in 1977 gives credence to this belief.

Until the first human influenza virus was isolated in 1933, it was impossible to tell with certainty which pandemics of severe respiratory disease were caused by influenza viruses. Since 1933, major antigenic shifts occurred in 1957 when the H2N2 subtype (Asian influenza) replaced the H1N1 subtype, in 1968 when the Hong Kong (H3N2) virus appeared, and in 1977 when the H1N1 virus reappeared.

III. Nomenclature

Influenza viruses are divided into Types A, B, and C, based on antigenic differences between their nucleoprotein (NP) and matrix (M) antigens. Influenza A viruses are further subdivided into subtypes and the revised nomenclature system (WHO, 1980) includes the host of origin, geographical origin, strain number, and year of isolation (Table 1). The antigenic description of the hemagglutinin and neuraminidase is given in parenthesis, *e.g.*, A/Swine/Iowa/15/30 (H1N1). By tradition,

Table 1. *Hemagglutinin and neuraminidase subtypes of influenza A viruses isolated from humans, lower mammals, and birds*

Subtype	Species of origin[c]			
	Humans	Swine	Horses	Birds
Hemagglutinin				
H1[a] (H0, H1, Hswl)[b]	PR/8/34	Sw/Ia/15/30	—	Dk/Alb/35/76
H2 (H2)	Sing/1/57	—	—	Dk/Ger/1215/73
H3 (H3, Hav7, Heq2)	HK/1/68	Sw/Taiwan/70	Eq/Miami/1/63	Dk/Ukr/1/63
H4 (Hav4)	—	—	—	Dk/Cz/56
H5 (Hav5)	—	—	—	Tern/S.A./61
H6 (Hav6)	—	—	—	Ty/Mass/3740/65
H7 (Heql, Havl)	—	—	Eq/Prague/1/56	FPV/Dutch/27
H8 (Hav8)	—	—	—	Ty/Ont/6118/68
H9 (Hav9)	—	—	—	Ty/Wis/1/66
H10 (Hav2)	—	—	—	Ck/Ger/N/49
H11 (Hav3)	—	—	—	Dk/Eng/56
H12 (Havl0)	—	—	—	Dk/Alb/60/76
H13	—	—	—	Gull/Md/704/77
Neuraminidase				
N1 (N1)	PR/8/34	Sw/Ia/15/30	—	Ck/Scot/59
N2 (N2)	Sing/1/57	Sw/Taiwan/70	—	Ty/Mass/3740/65
N3 (Nav2–3)	—	—	—	Tern/S.A./61
N4 (Nav4)	—	—	—	Ty/Ont/6118/68
N5 (Nav5)	—	—	—	Sh/Austral/1/72
N6 (Navl)	—	—	—	Dk/Cz/56
N7 (Neql)	—	—	Eq/Prague/1/56	FPV/Dutch/27
N8 (Neq2)	—	—	Eq/Miami/1/63	Dk/Ukr/1/63
N9 (Nav6)	—	—	—	Dk/Mem/546/74

[a] Current subtype designation (WHO 1980).
[b] Previous subtype designation (WHO 1971).
[c] The reference strains of influenza viruses, or the first isolates from that species are presented.

the host of origin of human strains is not included, *e.g.*, A/PR/8/34 (H1N1). Immunological, hybridization, and sequence data have indicated that several subtypes of the previous nomenclature (WHO, 1971) are so similar that they are now merged, *e.g.*, H0, H1, and Hsw1 became H1 (Table 1). Earlier designations indicating the species of origin of the HA and NA antigens are now omitted.

Since 1971, there have been extensive studies on influenza viruses in animals and there have been a few additions to the number of hemagglutinin and neuraminidase subtypes. These additions include three avian strains with new hemagglutinins (H9, H12, and H13, Hinshaw *et al.*, 1982) and two avian strains with new neuraminidase subtypes, N5 and N9. Extensive surveillance has not revealed further subtypes, suggesting there may be a finite number of distinct influenza A viruses in nature and that representatives of each of these subtypes continue to circulate primarily in lower animals.

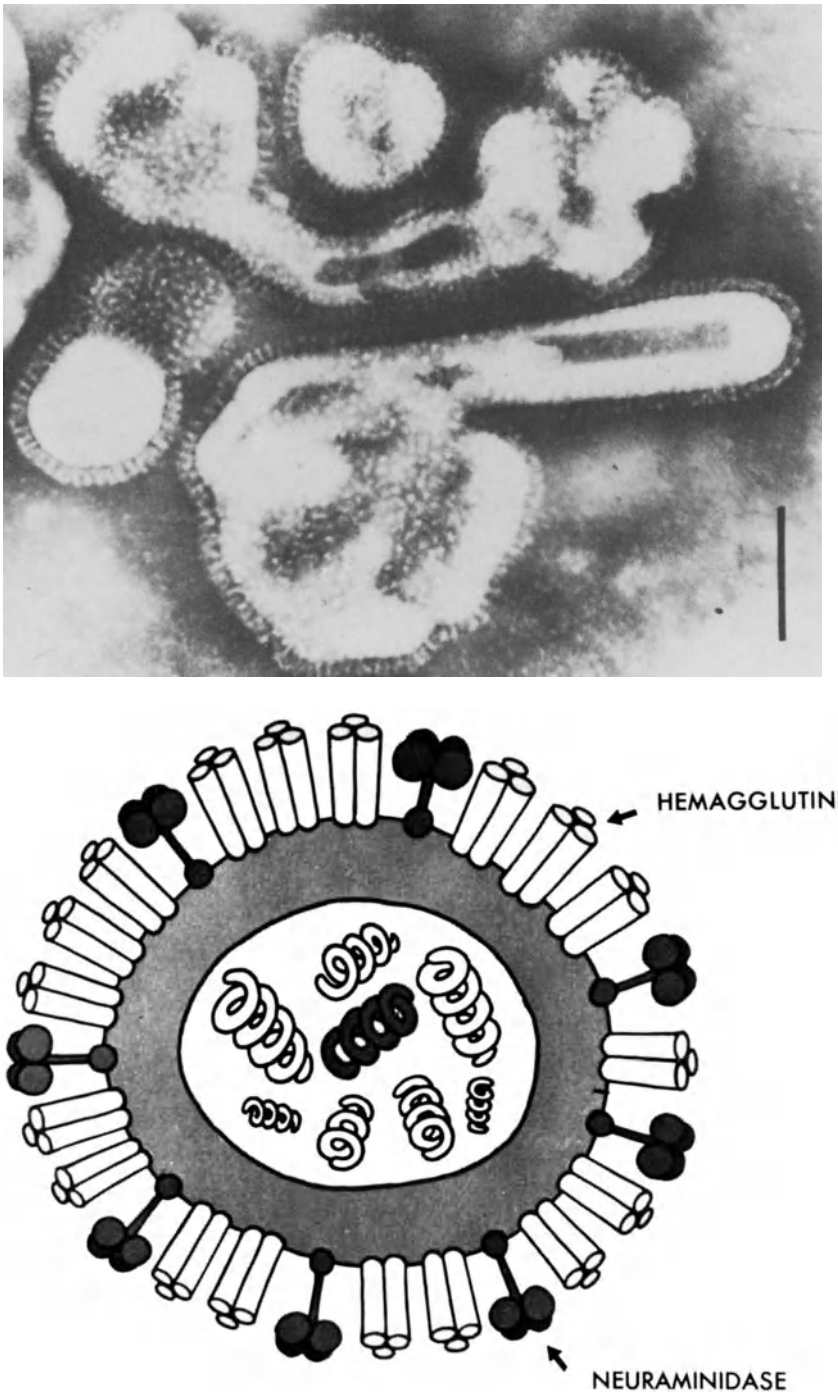


Fig. 1. Electron micrograph (top) and diagram showing the structure of particles of influenza type A virus. The hemagglutinin spikes are approximately 16 nm in length. The coiled structures inside the virus represent the 8 RNA segments of the virus, each of which codes for one (in some cases two, using overlapping reading frames) of the virus proteins.

IV. Chemical and Physical Properties of the Antigens

Particles of influenza A viruses are covered by a layer of "spikes" which are the two surface antigens of the virus, the hemagglutinin (HA) and the neuraminidase (NA) (Fig. 1). The HA is a trimeric rod-shaped molecule while the NA is a tetramer consisting of a square, box-like head on a long thin stalk. Both antigens are glycoproteins and they are attached to the lipid envelope of the virus by sequences of hydrophobic amino acids which occur at the base of the HA trimer and NA stalk. The HA and NA comprise about 25% and 5%, respectively, of the total protein in the virion.

Within the lipid envelope of the virus lies the M (membrane or matrix) protein, which is the major protein component of the virus and is believed to be structural in function. Within the matrix shell are eight single-stranded RNA molecules of negative sense (that is, the virion RNA is complementary to the messenger RNA) associated with a nucleocapsid protein and three large proteins, P1, P2, and P3, responsible for RNA replication and transcription. At least three virus-encoded nonstructural proteins (NS_1, NS_2, and M_2) are found in infected cells, but their functions are unknown. Chapter 2 gives the current information on the eight genes of the influenza virus and the molecular weights of the proteins encoded by them.

V. The Hemagglutinin (HA)

The HA accounts for about 25% of the virus protein. It is responsible for the attachment of virus to cells and for the penetration of the virus into the cell during the initial stages of infection. Antibodies to the HA neutralize the infectivity of the virus and it is variation in the HA glycoprotein which is mainly responsible for the continually occurring outbreaks of influenza and for our inability to control these by vaccination.

The HA monomer is coded by the fourth largest RNA segment and is synthesized as a single polypeptide chain which undergoes post-translational cleavage in at least three places (Fig. 2). An N-terminal signal sequence is removed and, depending on the host cell and virus strain, the molecule is cleaved, with the removal of one or more intervening residues, to give two polypeptide chains called HA1 and HA2, with molecular weights of 36,000 and 27,000, respectively (Chapter 2). HA1 and HA2 remain joined by a single disulphide bond and each HA "spike" contains three of these HA1 and HA2 chains. Cleavage of the HA polypeptide into HA1 and HA2 is necessary for the virus particles to be infectious. Virus grown in cells which lack the cleavage enzyme has low infectivity but this can be regained by treating the virus with trypsin. A sequence of 25—32 hydrophobic amino acids near the C terminus of HA2 serves to anchor the HA in the virus membrane.

A. Isolation and Antigenic Properties of HA

Intact HA "spikes" can be isolated from some strains of influenza virus following disruption of virus particles with a detergent, such as sodium dodecyl sulphate

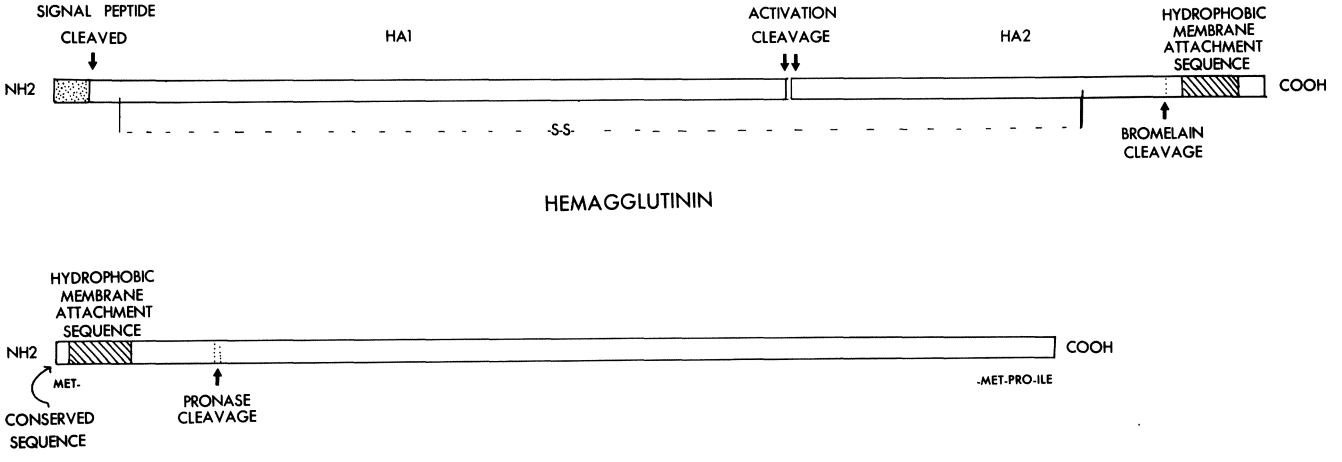


Fig. 2. Diagramatic representation of certain features of the hemagglutinin and neuraminidase polypeptides

1. Hemagglutinin. The HA is synthesized as a single polypeptide. Following its synthesis an N-terminal signal peptide is cleaved off and the molecule is cleaved further into HA1 and HA2 with the removal of one or more intervening amino acids. This latter cleavage is necessary for the virus to be infectious. HA1 and HA2 remain linked by a single disulphide bond and each HA spike contains three of these dimers. A sequence of hydrophobic amino acids near the C-terminus of HA2 serves to anchor the HA in the lipid of the virus membrane. Treatment with bromelain removes this hydrophobic region without damaging the rest of the molecule which, in some cases, can be crystallized

2. Neuraminidase. The neuraminidase is oriented in the virus membrane in the opposite way to the hemagglutinin. No post-translational cleavage of the NA polypeptide occurs, no signal peptide is split off and even the initiating methionine is retained. No processing at the C terminus takes place—the C-terminal sequence, -MET-PRO-ILE predicted from the gene sequence is found in intact NA molecules isolated from virus and in the pronase-released NA heads. A sequence of six polar amino acids at the N terminus of the NA polypeptide, which is totally conserved in at least 8 different NA subtypes, is followed by a sequence of hydrophobic amino acids which probably represents the transmembrane region of the NA stalk. This sequence is not conserved at all between subtypes (apart from conservation of hydrophobicity). Pronase cleaves the polypeptide in the positions shown, removing the stalk and releasing the enzymatically and antigenically active head of the NA which, in some cases, can be crystallized

(SDS). When the detergent is removed from purified preparations of HA, the HA molecules aggregate by their hydrophobic ends to form rosettes (Fig. 3). When these rosettes are mixed with purified IgG antibody to the HA, the antibody molecules can be seen in electron micrographs to bind to the hydrophilic end of the HA, just below its tip (Fig. 3).

Incomplete HA molecules can be isolated from some strains of influenza virus by digesting the virus particle with proteolytic enzymes such as bromelain or pronase. In Hong Kong (H3N2) virus, bromelain cleaves the HA between residues 175 and 176 in HA2, removing the hydrophobic membrane attachment sequence and releasing the hydrophilic end of the HA. Antibody molecules can be seen to bind to the bromelain-released HA near one of the ends (Fig. 3), but in this case it is not possible to tell which end of the molecule is which.

Bromelain-released HA from some Hong Kong (H3N2) viruses can be crystallized and from x-ray diffraction data the three-dimensional structure of Hong Kong (H3N2) HA has been determined (Wilson *et al.*, 1981; Wiley *et al.*, 1981). Conservation of structural features such as disulphide bonds suggests that HA molecules from other subtypes have similar structures. The HA glycoprotein of influenza virus is a trimer built of two structurally distinct regions: a triple-stranded coiled-coil of α-helices which extends 76 Å from the membrane and a globular region of antiparallel β-sheet that contains the receptor binding site. The variable antigenic determinants are located on the top domain. Each subunit has an unusual loop-like topology: it begins at the membrane, extends 135 Å distally and folds back to enter the membrane (Fig. 4). Early experiments showed that two groups of antigenic determinants, strain-specific and common (cross-reacting), existed in the HA molecule and the electron micrographs showed that these were apparently located at the side of the HA spike, near its tip. No further dissection of the antigenic structure of the HA was possible using polyclonal antisera.

In addition to the virus-specified antigenic sites on the HA, the HA carries host-cell specific antigenic sites which are carbohydrate in nature (Ward, 1981). A/Mem/72 (H3N2) HA possesses 7 oligosaccharide side chains, three of which (at ASN 8 and ASN 22 in HA1 and ASN 154 in HA2) are antigenically related to chick embryo host antigen. Jap/305/57 (H2N2) HA possesses 5 oligosaccharide side chains of which only those at ASN 11 in HA1 and ASN 159 in HA2 are related to host antigen.

B. Changes in Conformation and Antigenicity of the HA at Low pH

Examination of the amino acids of the amino-terminal sequence of HA2 of several influenza A viruses of human and avian origin has revealed a highly conserved region of 10 residues (Skehel and Waterfield, 1977; Scheid and Choppin, 1977). This sequence is homologous to the amino-terminus of the cleaved, activated fusion glycoprotein (F) of the paramyxovirus, Sendai virus. The F protein mediates fusion of the Sendai virus envelope with the plasma membrane of host cells, a funtion which is considered essential for virus infectivity. It has been postulated, therefore, that the infectivity of influenza virus is dependent on a fusion function of HA analogous to that of the Sendai fusion protein. Although the three-dimensional

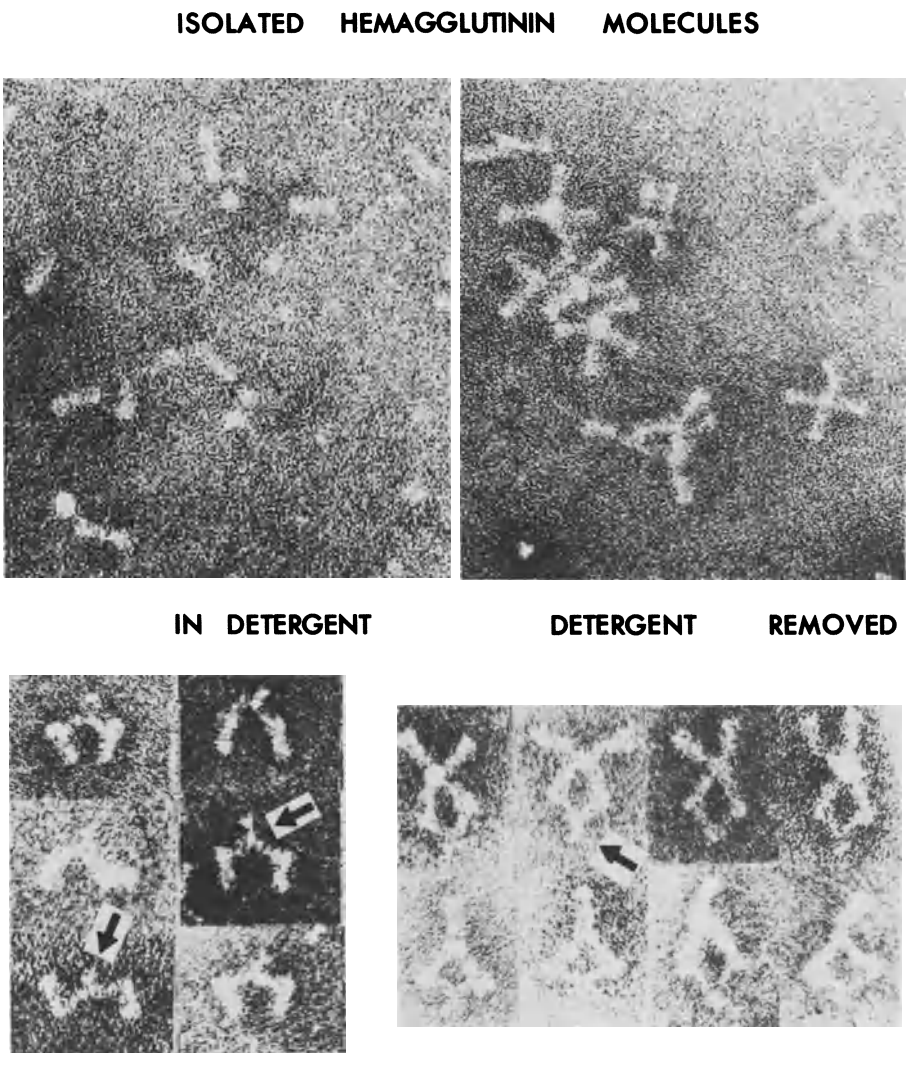


Fig. 3. Electron micrographs showing isolated hemagglutinin molecules. The top pictures show HA isolated after disruption of virus with sodium dodecyl sulphate (SDS). The HA exists as individual molecules in the presence of the detergent (left) which aggregate by their hydrophobic ends (these serve to attach the HA to the lipid of the virus envelope) when the detergent is removed (right). The lower pictures show the attachment of IGG antibody molecules to the HA. On the left is bromelain-released HA which has lost its hydrophobic membrane-attachment region and exists as individual molecules in the absence of detergent. On the right is aggregated, detergent-released intact HA. Anti-HA IGG antibody molecules (indicated by arrows) can be seen attached to the HA molecules. The antibody seems to bind to the HA just below the tip of the HA spike. Electron micrographs were taken by Robin Valentine and Nick Wrigley

Table 2. *Antigenic Drift of H3N2 viruses between 1968 and 1977 as demonstrated by their cross-reactions in hemagglutination-inhibition tests (from Pereira, 1979)*

	Ferret antiserum							
Virus	A/HK/1/68	A/E/42/72	A/PC/1/73	A/Sc/840/74	A/Vic/1/75	A/Tok/1/75	A/E/864/75	A/Tex/1/77
A/Hong Kong/1/68	2560	2560	320	80	20	<20	20	20
A/England/42/72	40	1280	320	160	40	<20	40	20
A/Port Chalmers/1/73	20	640	640	160	40	<20	40	20
A/Scotland/840/74	<20	40	40	640	<20	<20	<20	<20
A/Victoria/1/75	<20	<20	<20	20	1280	80	160	20
A/Tokyo/1/75	<20	<20	<20	<20	320	1280	40	40
A/England/864/75	<20	80	80	20	320	<20	5120	640
A/Texas/1/77	<20	40	80	20	160	<20	5120	1280

structure of the influenza HA shows that the amino terminus of HA2 is buried in the HA molecule, it has recently been shown that after incubation in conditions favoring membrane fusion, that is, at pH 5.0, a conformational change occurs in the HA glycoprotein (Skehel *et al.*, 1982). The molecule acquires the ability to hemolyse erythrocytes (Maeda and Ohnishi, 1980; Huang *et al.*, 1981) and bromelain-released HA can bind to lipid vesicles. The change in conformation is accompanied by a change in antigenicity. Monoclonal antibodies prepared to pH 5-treated hemagglutinin detect new antigenic determinants at this pH and the antigenic sites B and D (Fig. 4) present on the molecule at neutral pH are lost or modified at pH-5 (Webster *et al.*, 1983).

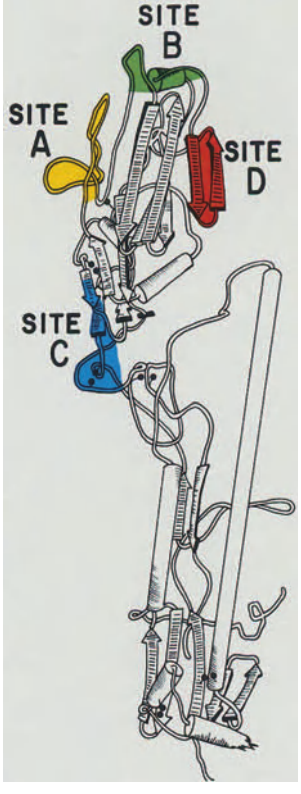


Fig. 4. Drawing (by Hidde Pleogh) of the Hong Kong HA monomer showing folding of the HA1 and HA2 polypeptides, Wilson *et. al.* (1981), Wiley *et. al.* (1981). The four colored areas show where four independent antigenic areas may be located. Note that in the HA spike, which is composed of three of the monomers, the red site is buried and may not be involved in antibody binding

C. Antigenic Drift in the HA

Antigenic drift occurs in the HA of type A influenza virus following the emergence of a new subtype. This has been well documented in the Hong Kong (H3N2) series of influenza viruses and is shown in Table 2.

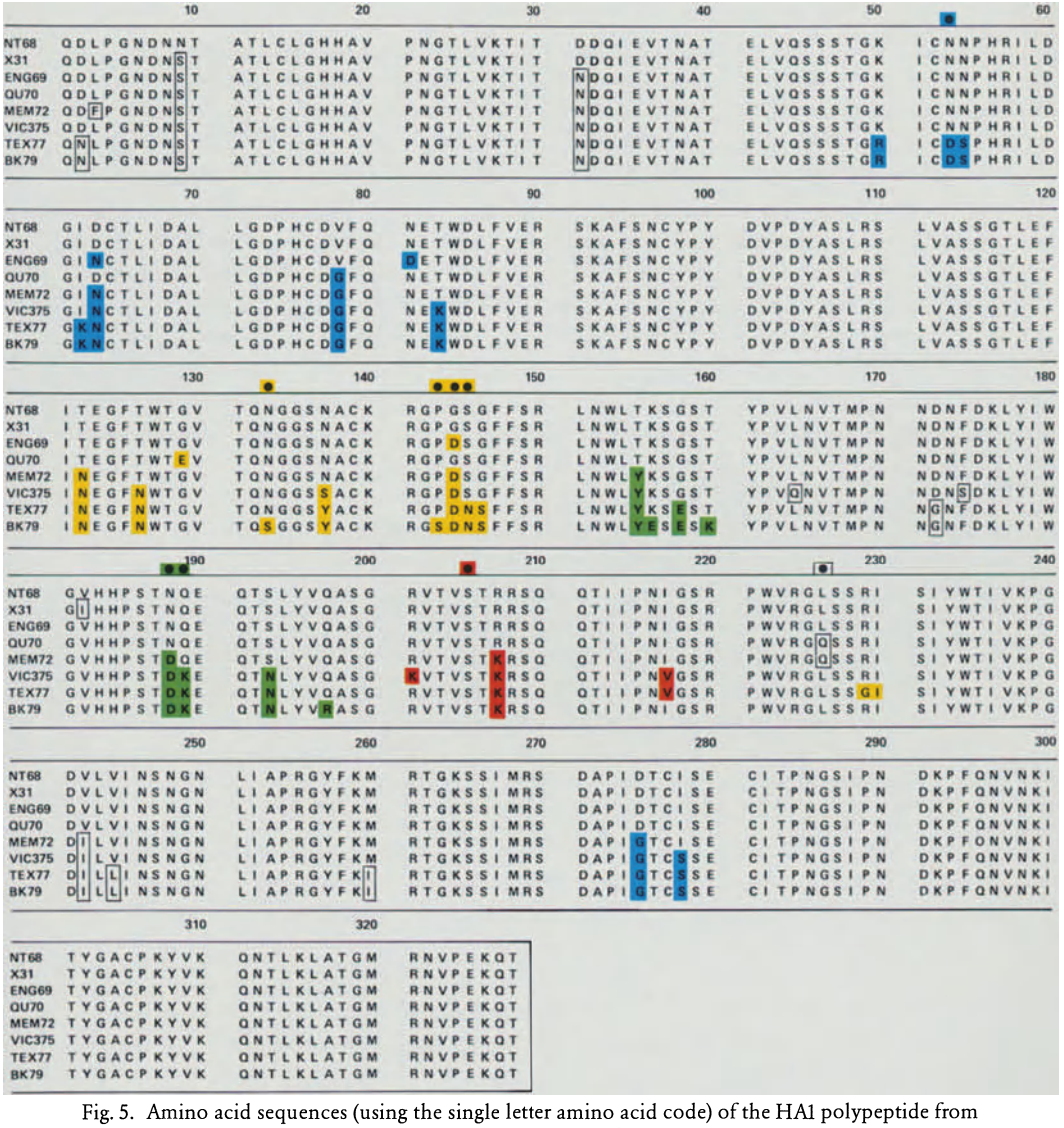


Fig. 5. Amino acid sequences (using the single letter amino acid code) of the HA1 polypeptide from eight variants of Hong Kong (H3N2) influenza virus isolated between 1968 and 1979. Residues which have changed from NT68 are boxed. The location of most of these changes on the three-dimensional structure of the HA are shown in Fig. 6. Dots indicate the single sequence changes found in variants selected with monoclonal antibodies. These are (53) ASN → LYS, (133) ASN → LYS, (143) PRO → SER, THR, LEU, HIS, (144) GLY → ASP, (145) SER → LYS, (205) SER → TYR, (188) ASN → ASP, (189) GLN → HIS. The change (226) LEU → GLN appears to arise with remarkable ease. This change has been found in different clones of the same field strain and in monoclonal variants which already show another change and may be unrelated to changes in antigenicity

Antigenic drift in the HA occurs by successive mutations in the gene, leading to an accumulation of amino acid sequence changes that alter the antigenic sites in such a way that they are no longer recognized by the host's immune system. Evidence for this mechanism comes from sequence analysis of naturally occurring "drifted" viruses (mainly those of the Hong Kong [H3N2] virus) (Table 2) and of variants selected in the laboratory in the presence of antibodies. Most of the sequence changes occurring in Hong Kong HA between 1968 and 1979 were in the HA1 portion of the HA molecule, only three changes being found in HA2 (Both and Sleigh, 1981). The sequence changes in HA1 during this period are shown in Fig. 5 in chronological order of isolation of the strains. Although these changes are distributed over the sequence of HA1, on the three-dimensional structure of the HA many are clustered into four regions (Fig. 4) which are assumed to be the antigenic sites on the HA. Sequential changes at a particular position in HA1 have not occurred. A few apparent reversions were seen, but as there is no reason to believe that the viruses examined represent a direct genealogical lineage, it is likely that in those cases in Fig. 5 where there are apparent reversions of an amino acid, the real explanation is that the residue never changed in the first place. Similarly, in the one case where there is an apparent sequential change in an amino acid (the asparagine residue at position 137 changing to serine in Vic/75 and then to tyrosine in Bangkok/1/79), the nucleotide changes involved indicate that the Vic/75 strain was not the progenitor of the Bangkok/1/79 strain.

Some regions of HA1 have been totally conserved during the period 1968–1979. The largest of these conserved regions occurs between residues 84–121 and 279–328. The latter block is also conserved in the two avian H3 strains so far sequenced (Fang *et al.*, 1981; Ward and Dopheide, 1982), except for a single change of valine to isoleucine at position 309. Residues 279–328 form part of the "stalk" of the HA, where structural constraints may necessitate that they are highly conserved.

The sequence changes which occurred in the HA1 portion of the HA of Hong Kong antigenic variants (field strains) isolated between 1968 and 1979 are shown in Fig. 5.

It is not known, however, which of these sequence changes are responsible for the changes in antigenicity which occurred in the HA during this period and which of the changes exemplify "random drift", having no effect on the antigenic properties of the HA.

D. Use of Monoclonal Antibodies in the Analysis of Antigenic Drift

Antigenic variants of the HA have been selected by growing virus in the presence of monoclonal antibody to the HA. The variants were present at a frequency of $\sim 10^{-5}$ in stock virus (Yewdell *et al.*, 1979; Webster *et al.*, 1980) and did not bind at all to the antibody used for their selection. Such variants have been used to construct an operational antigenic map of the hemagglutinin molecule by comparative antigenic analysis of the mutant viruses with the monoclonal antibodies.

Analysis of A/PR/8/34 (H1N1) variants by radioimmunoassay (RIA) revealed 4 antigenic sites on the HA molecule (Gerhard *et al.*, 1981). Two of these are

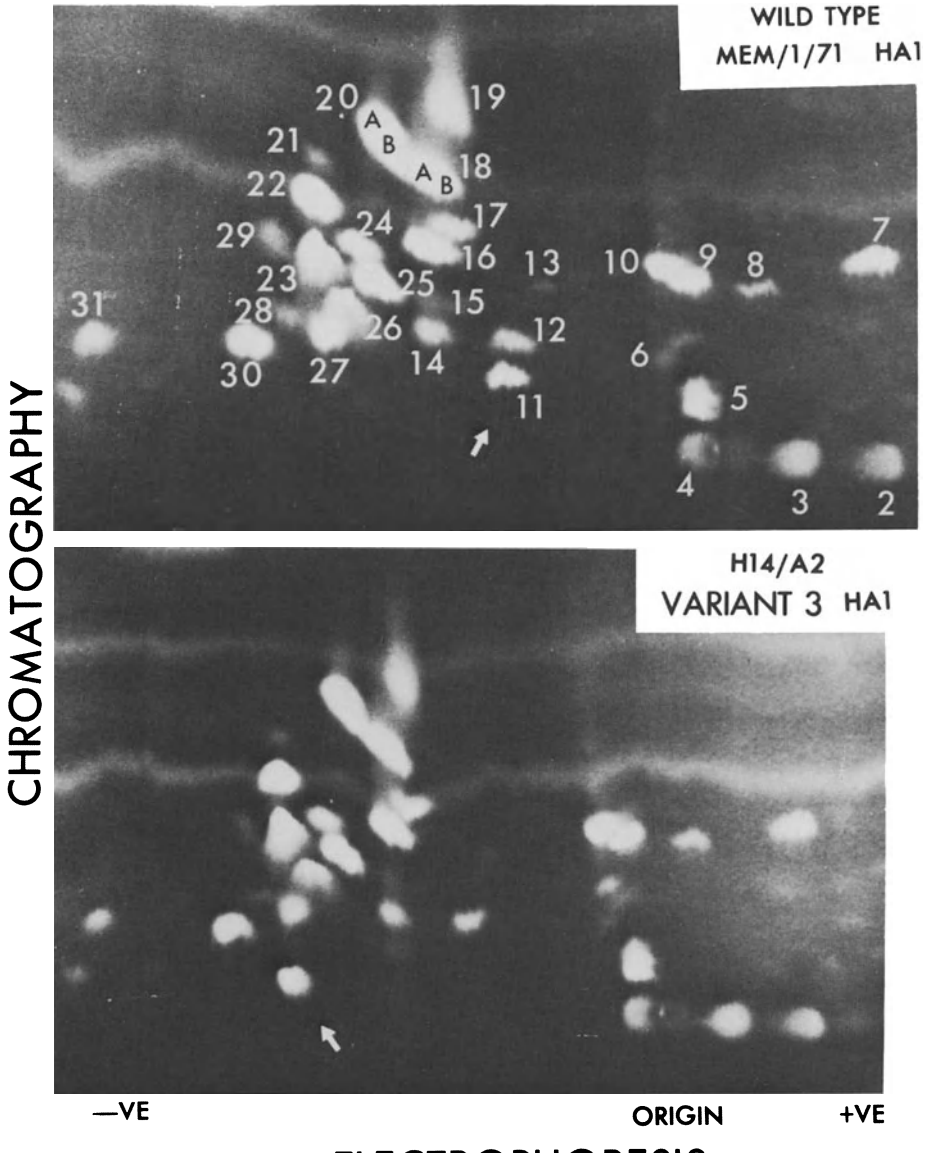


Fig. 6. Maps of the trypic peptides (soluble at pH 6.5) from S-carboxymethylated HA1 of wild-type Mem/71 virus and one of the four variants selected with H14/A2 monoclonal hybridoma antibody. The maps were stained with fluorescamine. A single peptide differences, peptide 11 (arrow), was seen on the maps. Peptide 11 comprised residues 51—57 (-ILE-LYS-ASN-ASN-PRO-HIS-ARG-) in HA1 from wild-type Mem/71 virus. The asparagine residue at position 53 was replaced by lysine in the variant. No differences were found in any of the other peptides

predominantly "strain-specific", designated Sa and Sb, and two are "cross-reactive", designated Ca and Cb (Russ *et al.*, 1981). On further analysis, the Ca area has been subdivided into two areas, Ca1 and Ca2, which can be discriminated by some monoclonal antibodies but not by others.

Analysis of A/Mem/1/71 (H3N2) variants using HI assays revealed at least 3 nonoverlapping antigenic regions (Webster and Laver, 1980).

E. Sequence Change in the HA of Influenza Virus Variants Selected with Monoclonal Antibodies

Sequence changes in the HA of influenza virus variants selected with monoclonal antibodies have been determined in the following way: HA was isolated from SDS-disrupted parental virus and HA1 and HA2 were separated, S-carboxymethylated and digested with trypsin. The tryptic peptides were separated by two-dimensional electrophoresis and chromatography and each was analyzed for amino acid composition (Fig. 6). The composition data allowed the peptides to be located in the HA1 and HA2 sequences and the sequence of each could be deduced in this way. Not all of the tryptic peptides could be analyzed, however. Two insoluble tryptic peptides in HA1 and several in HA2 could not be examined. Sequences of these regions have been determined by sequencing the RNA genome of the variant viruses (Newton and Air, unpublished results). In the case of some

Table 3. *Sequence changes found in the HA1 polypeptide from antigenic variants of A/Memphis/1/71 virus selected with the monoclonal antibodies listed*

Monoclonal antibody	Variant	Sequence change[a]
H14/A2	V1	Asparagine (53) → Lysine
	V2	Asparagine (53) → Lysine
	V3	Asparagine (53) → Lysine
	V4	Asparagine (53) → Lysine
H14/A21	V1	Serine (205) → Tyrosine
H14/A20	V1	Asparagine (133) → Lysine
	V2	Proline (143) → Serine
	V3	Proline (143) → Leucine
	V4	Serine (145) → Lysine
Mem/212/1	V1	Proline (143) → Serine
	V2	Proline (143) → Threonine
	V3	Proline (143) → Leucine
	V7	Proline (143) → Histidine
Mem/27/2	V5	Proline (143) → Serine
	V9	Proline (143) → Threonine
Mem/123/4	V1	Proline (143) → Histidine
	V3	Proline (143) → Histidine
	V10	Glycine (144) → Aspartic Acid
HK 30/2	V10	Asparagine (188) → Aspartic Acid

[a] Numbers in parentheses give the position of the amino acid in HA1.

variants, one or other of the tryptic peptides showed a quite striking difference in either electrophoretic or chromatographic mobility.

Fig. 6 shows an example of this. In this case the relevant peptide from the parental virus had the composition: ILE(1) ASP(2) CM-CYS(1) PRO(1) HIS(1) ARG(1) and in the peptide from the HA of the variant selected with H14/A2 monoclonal antibody, an ASP had been replaced by LYS.

This peptide in the parental virus occupied positions 51 to 57 in HA1 and had the sequence ILE-CYS-ASN-ASN-PRO-HIS-ARG, but from the composition it was not possible to tell which of the two asparagine residues had changed to lysine. The region of the RNA segment in the variant coding this peptide was therefore sequenced and it was determined that the ASN at position 53 changed to LYS. (The proximity of the CM-CYS residue apparently prevented tryptic cleavage at the new LYS residue.) Compositions of the other peptides in HA1 and HA2 were the same in the parental and variant viruses.

Similar analysis of the HA from a large number of variants of A/Memphis/1/71 virus selected with different monoclonal antibodies showed that in each case a single change in the amino acid sequence of HA1 occurred (Table 3, Fig. 5). No changes were found in HA2.

Since these sequence changes totally abolished binding of antibody which neutralized the infectivity of the parental virus, they clearly occurred in regions of the HA involved in the antigenicity of the molecule.

F. Location of the Antigenic Sites in the 3-D Structure of the HA

Many of the sequence changes occurring in HA1 from the natural isolates (field strains) of Hong Kong (H3N2) influenza virus isolated between 1968 and 1979 fell into 4 distinct regions when the linear HA1 polypeptide folded into its three-dimensional structure (Fig. 4).

The changes occurring in HA1 from the variants selected with monoclonal antibodies also fell into these four regions indicating that these are almost certainly the four nonoverlapping antigenic sites delineated by monoclonal antibodies. Most of the antigenic variants of Mem/1/71 (H3N2) selected with monoclonal antibodies could not be distinguished from parental virus with antisera from ferrets suggesting that they would have no epidemiological advantage in nature (Webster and Laver, 1980). A small number of variants selected with monoclonal antibodies to A/Mem/1/71 (H3N2), A/USSR/90/77 (H1N1) and B/Hong Kong/1/73 could be distinguished with ferret antisera (Kendal *et al.*, 1981; Webster and Berton, 1981). Presumably the amino acid substitutes in these variants caused changes in the antigenicity of the HA molecule. "Monoclonal" variants of A/Mem/1/71 (H3N2) with epidemiological potential could be reproducibly selected by saturating one antigenic region (Site A, Figure 4) with mutations before selecting variants in a second region (Webster *et al.*, 1981). None of the accumulated mutations in the first antigenic region produced a variant that could be discriminated from parental virus with ferret antiserum but all variants subsequently selected with monoclonal antibodies to another antigenic region were significantly different with ferret antisera.

Table 4. *Amino acid substitutions seen in HA1 of PR/8/34 variants selected with monoclonal antibodies of the five reactivity groups*

Sa				Sb				Ca_1				Ca_2				Cb			
	Amino acid				Amino acid				Amino acid				Amino acid				Amino acid		
Virus	mut[n]	position		Virus	mut[n]	position		Virus	mut[n]	position		Virus	mut[n]	position		Virus	mut[n]	position	
		H1	H3			H1	H3			H1	H3			H1	H3			H1	H3
CV1	K-E	161	165	BV1	E-G	152	156	SV3	G-R	236	240	DV4	S-P	136	140	AV1	R-G	74	82
2	P-L	158	162	2	Q-R	192	196	3	S-T	270	273	5	R-I	220	224	LV1	R-G	74	82
5	P-S	124	128	6	G-D	155	159	6	G-R	236	240	NV2	G-R	139	143	4	R-G	74	82
6	N-K	162	166	11	Q-L	188	192					7	S-A	141	145	7	S-P	75	83
JV1	S-F	163	167	12	N-S	189	193	WV8	I-M	178	182	TV1	G-R	139	143	MV4	L-P	71	79
2	K-E	161	165	13	N-K	189	193	8	S-L	203	207	26	D-G	221	225	5	S-P	75	83
3	S-Y	163	167	EV2	Q-R	188	192	10	V-A	165	169					9	V-E	73	81
9	S-F	163	167	6	Q-R	188	192									RV6	E-K	115	122
18	S-F	163	167	8	E-G	194	198	11	G-R	169	173					7	L-P	70	78
18	+K	161	165					15	G-E	236	240								
KV2	K-T	159	163					ZV1	G-V	169	173								
4	N-K	125	129																
PV9	S-L	156	160																
11	S-L	156	160																
12	N-Y	125	129																
20	E-V	154	158																

Data from Caton *et al.* (1982).
H1 numbering system is from Winter *et al.* (1981).
H3 numbering system is from Ward and Dopheide (1980) and Verhoeyen *et al.* (1980).

Caton *et al.*, 1982 have sequenced the HA1 coding portion of the HA genes of 46 variants of PR8 selected with monoclonal antibodies belonging to each of the 5 reactivity groups. The amino acid substitutions found are shown in Table 4. In two variants there were additional silent nucleotide changes, and more than one amino acid change was seen in 2 variants (JV 18 and WV 8).

The three-dimensional structure of the PR8 hemagglutinin is not known, but when the positions of amino acid variation are plotted on the drawing of the three-dimensional structure of the polypeptide chain of the Hong Kong (H3) HA monomer (Wilson *et al.*, 1981; Wiley *et al.*, 1981), clusters of variant amino acids are seen which in general correspond to the antigenic groupings. There are no variants of PR8 which map at the same residues as site "C" of Hong Kong HA, although the Cb site of PR8 is adjacent, while site "B" of Hong Kong HA becomes two sites in PR8. These changes may reflect differences in the three-dimensional structure between PR8 and Hong Kong hemagglutinins. Site Ca 1 of PR8 is especially interesting when plotted on the H3 map, since changes occur at three bends, each close to 180°, in the β-structure, in each case a small amino acid side chain changing to a larger one (GLY 173 to ARG or VAL, SER 207 to LEU, GLY 240 to ARG or GLU). In nature, the Sb and Sa sites underwent much more drift between 1933 and 1957 than did the Ca or Cb sites (Gerhard *et al.*, 1981). Antibodies directed against the Cb site had approximately 10-fold less virus-neutralizing potency than those of the Sa, Sb or Ca sites, which provides a possible explanation for lack of natural drift in the Cb area, but does not explain why the Ca site was relatively stable in nature (Gerhard *et al.*, 1981). It is possible that variants in the Ca site have an altered and disadvantageous conformational change in the HA.

G. Sequential Selection of Antigenic Variants

Variants selected with monoclonal antibodies had clearly lost an antigenic site on the HA since, in each case, the monoclonal antibody (which bound very efficiently to the wild-type HA) failed to bind at all to the variant.

Experiments were done to see if a new antigenic site was created on the hemagglutinin of the variants when the original site was lost. Hyperimmune antisera to the variant molecules were absorbed with purified concentrated wild-type virus until the hemagglutinin-inhibition titers of the sera for the wild-type virus reached undetectable levels. The absorbed sera were then tested against the variant viruses. When sera raised against the variants selected with H14/A2 and H14/A21 monoclonal antibodies, where the sequence changes were asparagine 53 to lysine and serine 205 to tyrosine, respectively (Laver *et al.*, 1979), were absorbed with wild-type Mem/71 virus, no hemagglutination-inhibition activity to the variants remained (Webster and Laver, unpublished results). This suggested that no new immunogenic site was formed in these variants. The same result was obtained whether rabbit or mouse antisera were used.

On the other hand, when serum raised against the variant Mem 123/4 (PRO 143 to HIS, Table 3) was absorbed with wild-type Mem/71 virus, high levels of inhibition activity to the variant remained after all inhibition activity to the wild-type

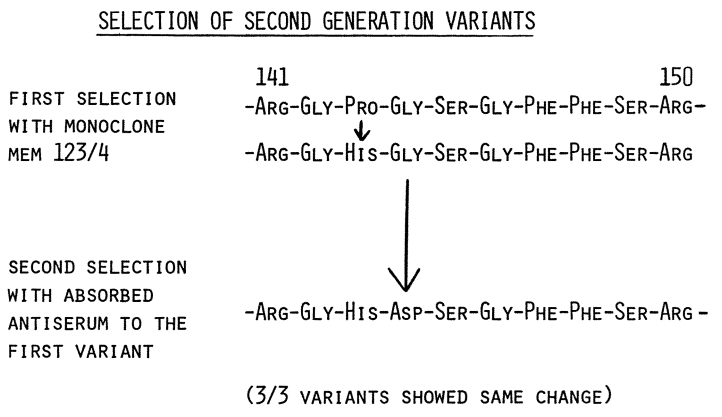


Fig. 7. Sequence changes during the sequential selection of antigenic variants of Mem/71 virus. In the first variant, selected in the presence of mouse monoclonal hybridoma antibody to Mem/71 HA, the proline at position 143 in HA1 changed to histidine. In the second selection, the initial variant was grown in the presence of antibody raised against the new antigenic site on this variant created by the sequence change of PRO (143) to HIS. This antibody was prepared by absorbing with wild-type Mem/71 virus, hyperimmune rabbit antiserum raised against the variant HA

virus had been removed. This antibody to the variant was not removed after repeated absorption of the sera with wild-type virus, and therefore must have been directed against a single new antigenic site on the variant hemagglutinin. In this respect it behaved like "monoclonal" antibody. Similar results were obtained when antibody to the hemagglutinin of the variant which showed a change at position 143 in HA1 of proline to threonine (Table 3) was absorbed with wild-type virus.

The antibody raised against the new antigenic site on the hemagglutinin molecule from the variant viruses was used to select second-generation variants of the initial variants. When the first sequence change in the variant was PRO (143) to HIS, the sequence change in the second selection was GLY (144) to ASP (Fig. 7). When the first sequence change in the variant was PRO (143) to THR, in the second selection the new THR at position 143 reverted to PRO and the virus regained the antigenic property of wild-type.

H. Antigenic Drift in the HA of Influenza A Virus from Lower Animals

Antigenic differences have been detected within the subtypes that include swine, equine, and avian influenza viruses, but the evidence points to cocirculation of different variants rather than to gradual antigenic drift of a particular virus. Studies with swine influenza viruses (H1N1) indicate that antigenic variation has occurred since the virus was first isolated in 1930 and that antigenically distinguishable viruses cocirculate (Kendal *et al.*, 1977). With swine influenza viruses there is

Table 5. *Antigenic comparison between equine influenza viruses (H3N8) from 1963–1981*

Virus	Hemagglutination-inhibition titers* with ferret antisera to:					
	Eq/Miami/1/63	Eq/Uruguay/1/63	Eq/Ky/1/76	Eq/Fontainebleu/1/79	Eq/Switz/2225/79	Eq/Ky/1/81
Eq/Miami/1/63	*160*	80	20	20	40	20
Eq/Uruguay/1/63	40	*160*	160	80	20	160
Eq/Ky/1/76	20	80	*160*	80	<20	40
Eq/Fontainebleu/1/79	40	160	320	*160*	20	80
Eq/Switz/2225/79	40	<20	<20	<20	80	<20
Eq/Ky/1/81	20	80	320	80	<20	*160*

* HI titer = reciprocal of serum dilution inhibiting four hemagglutinating doses of virus.

evidence for genuine antigenic drift with the appearance and disappearance of strains. Analysis of equine influenza viruses (H3N8) with ferret antisera and with monoclonal antibodies to the HA indicate that limited antigenic variation does occur (Pereira, 1969). At least two antigenically distinguishable equine 2 viruses continue to circulate (Hinshaw *et al.*, 1982). The A/Eq/Kentucky/1/81 (H3N8) strain that is currently causing disease problems in horses is antigenically similar to a virus isolated 18 years earlier (A/Eq/Uruguay/1/63 [H3N8]) whereas A/Eq/Switz/2225/79 (H3N8) is similar to A/Eq/Miami/1/63 (H3N8) (Table 5). Antigenic analyses of the HA of avian influenza viruses with ferret antisera (Hinshaw *et al.*, 1980) or with monoclonal antibodies (Hinshaw *et al.*, 1981) indicate that many different strains cocirculate, but only limited antigenic drift occurs. Serological analysis of influenza viruses from animals and birds therefore indicates that antigenic variation is less extensive than in human strains and that multiple variants cocirculate.

Comparison of the RNA sequences of influenza A virus HAs indicates that the differences within a single subtype of human or non-human strains are generally less than 10% (Fig. 11) (Air *et al.*, 1981). This suggests that the extent of mutations in human and animal strains is similar. Why then do equine and avian strains apparently show less antigenic variation in serological assays? The reason may be that to date only antigenically distinct cocirculating strains have been examined biochemically. The limited amount of antigenic variation found in equine and avian strains and the cocirculation of multiple variants of a subtype are features that distinguish these viruses from human strains. One possible explanation is that there is little immunological pressure to select variants in birds and in horses. Recent studies of horses by Schild and associates (personal communication) indicates that the equine antibody response is relatively short lived; circulating antibodies disappear within a few months. In avian species, the humoral immune response following infection is also short lived (Kida *et al.*, 1980) and each spring a large virgin population of susceptible animals becomes available. The puzzle about birds is the maintenance of multiple strains of an influenza virus subtype between annual outbreaks. Are they maintained by circulation in the population, are they preserved frozen in pond water, or are they integrated in some fashion in the host? Whatever the mechanism, it is efficient, because the multiple variants are maintained over many years.

VI. The Neuraminidase (NA)

The neuraminidase molecule is composed of a single polypeptide chain. It exists as a mushroom-shaped spike on the surface of the virion (Fig. 1) and may play a co-operative role with the hemagglutinin in promoting the fusion of viral envelope and cell membrane during the early stages of infection (Huang *et al.*, 1981). It has a box-like head made out of four coplanar and roughly spherical subunits and a centrally attached stalk containing a hydrophobic region by which it is embedded in the viral membrane (Fig. 8) (Laver and Valentine, 1969; Wrigley *et al.*, 1973; Wrigley, 1979). The emerging three-dimensional crystal structure (Colman and Laver, 1981) shows that the NA is a tetramer.

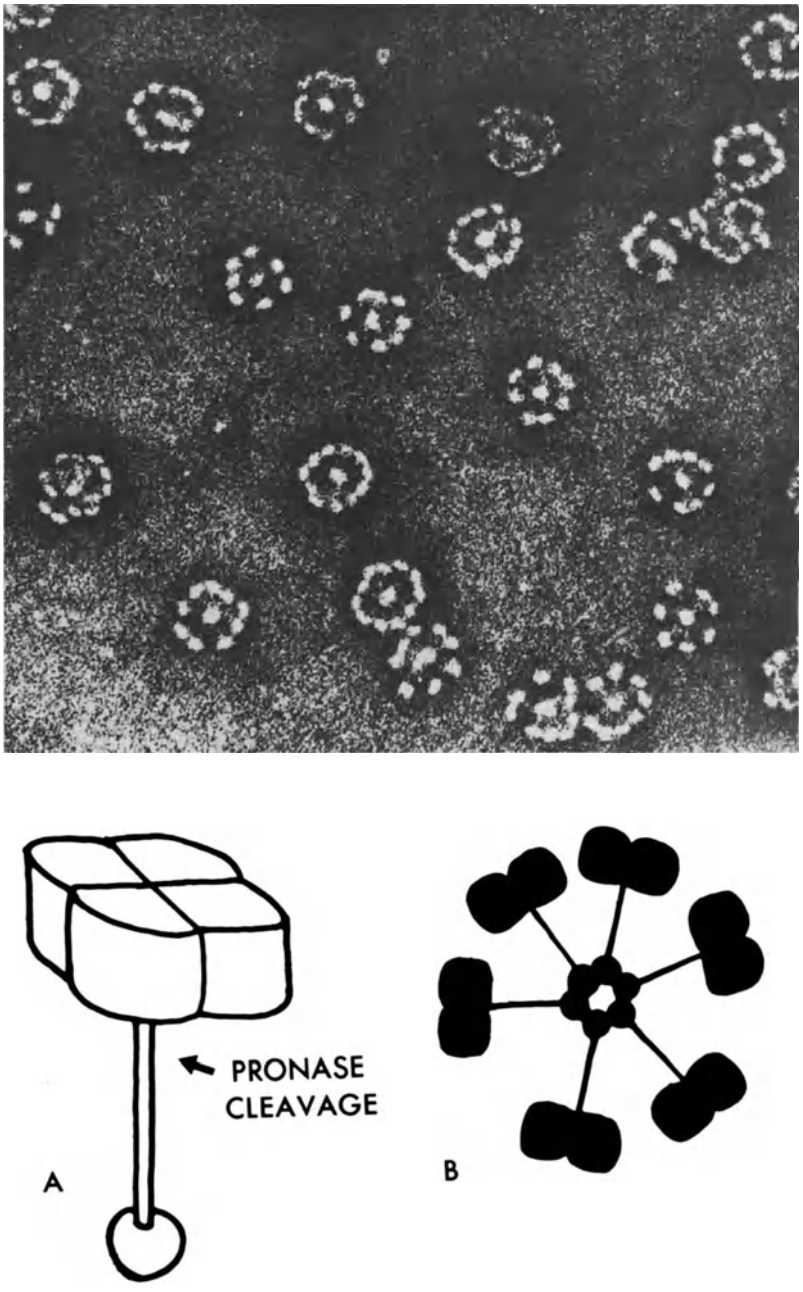


Fig. 8. Electron micrograph and diagram showing detergent-released neuraminidase molecules from which the detergent has been removed. The NA molecules have aggregated by the hydrophobic region near the end of the stalk which served to attach the NA to the lipid of the virus envelope. Treatment of virus particles with pronase releases the head of the NA which carries the enzymatic and antigenic acitivities of the molecule and which, in some cases, can be crystallized. Electron micrograph by Nick Wrigley

The neuraminidase is anchored in the virus membrane at its amino terminus, opposite to the Carboxy-terminal attachment of the hemagglutinin (Fig. 2). No post-translational cleavage of the NA polypeptide occurs; no signal peptide is split off and even the initiating methionine is retained. No processing at the C terminus takes place: the C-terminal sequence, -MET-PRO-ILE, predicted from the gene sequence is found in intact NA molecules isolated from virus, and in the pronase-released NA heads. A sequence of six polar amino acids at the N terminus of the NA polypeptide, which are totally conserved in at least 8 different NA subtypes, is followed by a sequence of hydrophobic amino acids which probably represent the transmembrane region of the NA stalk. This sequence is not conserved at all between subtypes (apart from conservation of hydrophobicity). Pronase cleaves the polypeptide in the positions shown (Fig. 2 and 8), removing the stalk and releasing the enzymatically and antigenically active head of the NA which, in some cases, can be crystallized (Fig. 10).

A. Antigenic Variation in the Neuraminidase

Antigenic shift and antigenic drift occur in the NA; among the human influenza viruses antigenic shift occurred in 1957, with the emergence of the H2N2 subtype and the replacement of N1 by N2. Antigenic drift in the NA of influenza viruses (Paniker, 1968; Schild and Newman, 1969; Curry *et al.*, 1974) has been correlated with differences in amino acid sequences (Laver, 1978; Kendal and Kiley, 1973).

Partial sequences up to 340 nucleotides from the 3′ end of the NA gene have been obtained for representative viruses from eight of the nine NA subtypes (Blok and Air, 1982 a). These sequences correspond to the region encoding the N-terminal part of the NA, since analysis of tryptic peptides shows that translation begins from the first AUG codon and that no signal peptide is cleaved from the N terminus (Blok and Air, 1982 a). The first six amino acids are conserved in all viruses of all subtypes examined and the next six are conserved in most subtypes. Following these, the nucleotide and predicted amino acid sequences of the eight subtypes of NA differ dramatically in the hydrophobic membrane insertion sequence and the extended structure of the stalk. Deletions and/or insertions of short blocks of nucleotides in the NA gene are seen within subtypes (Blok and Air 1982 b).

Gene sequences (and the deduced amino acid sequences) of the NA from four N2 strains (Udorn/72, NT/60/68, UK/75 and RI/5^{+}/57) and two N1 strains (PR/8/34 and WSN) have been obtained. The overall homology between the amino acid sequences of the N1 and N2 neuraminidase was estimated to be 39%.

Pronase-released neuraminidase heads from six strains of influenza type A virus (of the H2N2 and H3N2 subtypes) isolated between 1957 and 1975 were examined for changes in amino acid sequence (Fig. 9). In 469 residues, 19 changes which occurred during this period were located. Two regions of the neuraminidase molecule were not examined, residues 188—210, which were insoluble, and residues 1—73 (or 76), which form part of the stalk of the neuraminidase and were removed during pronase digestion.

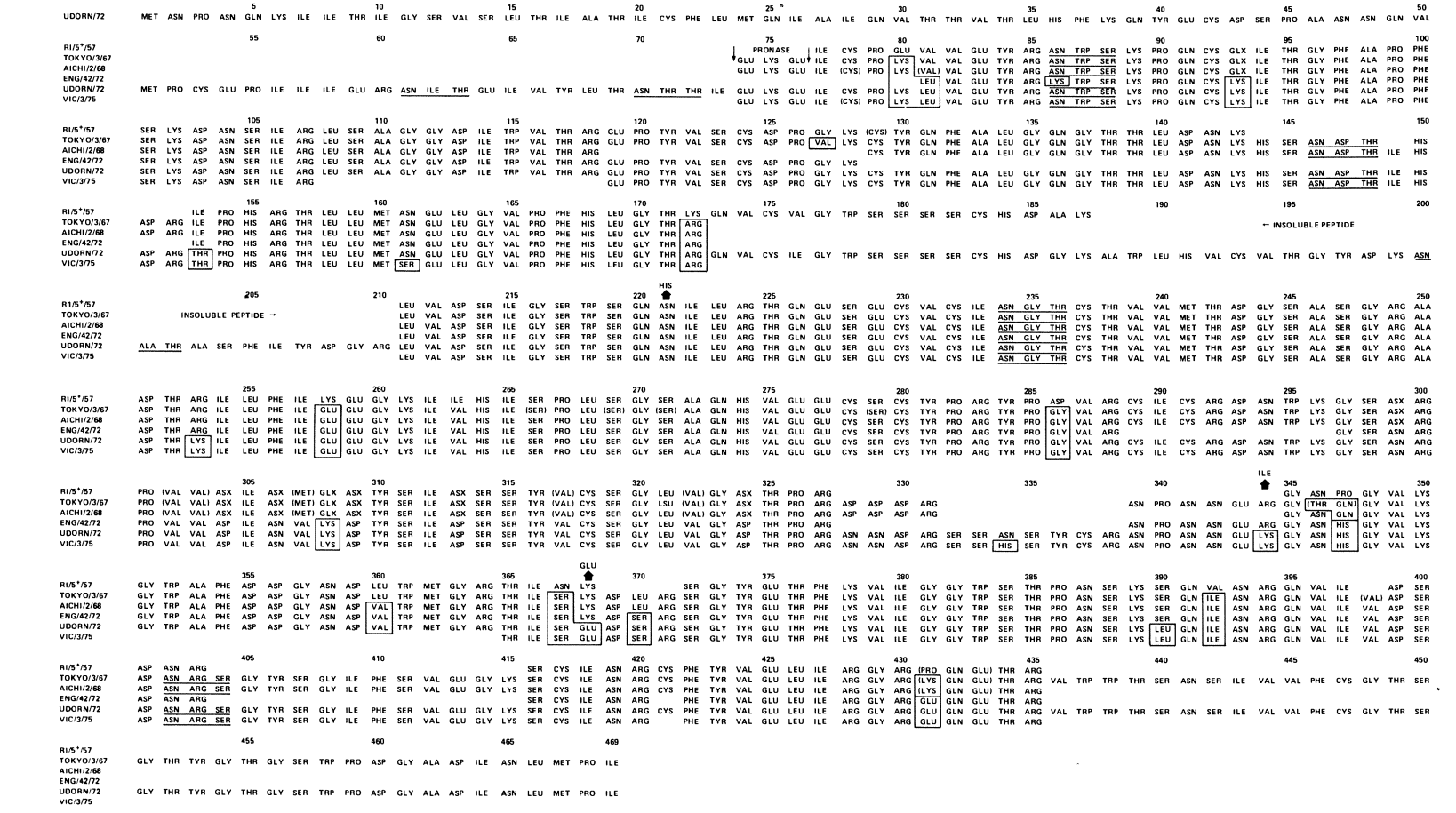


Fig. 9. Comparative partial amino acid sequences of the NA from 6 strains of N2 virus isolated between 1957 and 1975. The sequences werde deduced from the amino acid compositions of the tryptic peptides from pronase-released NA. Boxes indicate unambiguous differences in the sequences of the different NAs. Sequence changes of variants of A/Tokyo/3/67 virus selected with monoclonal antibodies are shown by arrows above the sequences. They are: (221) ASN → HIS, (368) LYS → GLU, (344) ARG → ILE (4 variants selected with different monoclonal antibodies showed this change)

The sequence changes were fairly evenly distributed along the NA polypeptide, but two obvious clusters could be seen. The residues in these regions may be involved in antigenic sites on the neuraminidase.

B. Selection of NA Variants with Monoclonal Antibodies

Antigenic variants of the NA of A/Tokyo/3/67 (H2N2) influenza virus have been selected by growing the virus for one passage in chick embryos in the presence of monoclonal antibodies to the NA. The variants were selected at a frequency of approximately 10^{-5} from a cloned virus (Webster *et al.*, 1982). The frequency of isolation of NA variants is similar to that for the HA (Yewdell *et al.*, 1979). A panel of different variants was selected and used to construct an operational antigenic map of the N2 molecule by comparative antigenic analysis of the mutant viruses with the monoclonal antibodies in ELISA and neuraminidase inhibition assays. These studies provided evidence for at least three nonoverlapping antigenic sites on the NA molecule. Further analysis using competitive RIA and enzyme inhibition of different sized substrates indicated that the antigenic sites are overlapping (Jackson and Webster, 1983), and that one antigenic site could be subdivided into two regions, making a total of four antigenic regions on N2 neuraminidase (Table 6). The monoclonal antibodies all inhibited enzyme-catalyzed hydrolysis of sialic acid from fetuin (MW 50,000) but only monoclonal antibodies in subgroup I inhibited enzyme activity on N-acetyl neuraminyl lactose (MW 600).

Neuraminidase heads from seven such variants of A/Tokyo/3/67 virus, selected with different monoclonal antibodies to the neuraminidase, were examined for sequence changes (Fig. 9). In four of these, a single sequence change at position 344 of arginine to isoleucine was found, in another variant an asparagine residue at position 221 changed to histidine and in another the lysine at position 368 changed to glutamic acid. This last change also occurred in the field strains isolated in 1972 and 1975. In the seventh monoclonal variant the change could not be found and may be in the insoluble region.

The four monoclonal variants with a change at residue 344 and the variant with a change at residue 368 were operationally grouped into a single antigenic region. X-ray crystallographic analysis will determine if these regions of the molecule are in close proximity in the three-dimensional structure.

Tokyo/67 NA has been crystallized and an electron density map at 3.1 Å has been calculated. The image shows large portions of the molecule to be in the β-sheet conformation and a complete chain tracing has recently been obtained (P. Colman, personal communication).

A Fab fragment from a monoclonal antibody to Tokyo/67 HA (S10/1) has also been crystallized (Fig. 10) and analysis of its structure is in progress. The S10/1 Fab crystals are trigonal, space group $P3_121$, with cell dimensions $a = 132.3$ Å, $c = 73.8$ Å. This crystalline Fab fragment forms a complex with the neuraminidase; the molecular weight of the complex is $406{,}000 \pm 20{,}000$ indicating that four Fab fragments are attached to each neuraminidase tetramer.

Table 6. *Amino acid substitutions in variants of A/Tokyo/3/67 (H2N2) and their antigenic grouping*

Monoclonal antibody to N2 of:		Antigenic variants[a]		Antibody groupings according to:		
	Rep #	Variant No.	Amino acid substitution	NI and[b] ELISA	Inhibition of enzyme activity on neuraminyl lactose	Competitive[c] RIA
Tokyo/3/67	25/4	V1	Arg $\rightarrow_{344}$ Ile	Ia	+	1
	S25/3	V1	Arg $\rightarrow_{344}$ Ile	Ia	+	1
	S32/3	V1	Arg $\rightarrow_{344}$ Ile	Ia	+	1
	23/9	V2	NT	Ib	+	1
	16/8	V1	NT	Ib	−	NT
	S10/1	V1	Lys $\rightarrow_{368}$ Gly	Ia	−	2
Japan/305/57	113/2	V1	NT	II	−	3
Texas/1/77	18/1	V1	Asn $\rightarrow_{221}$ His	III	NT	NT
	67/1	V1	NT	III	NT	NT

[a] Antigenic variants were selected with monoclonal antibodies (Webster *et al.*, 1982) and the amino acid substitutions were determined as described by Laver *et al.*, 1982.
[b] Reactivity in neuraminidase inhibition (NI) and ELISA as described by Webster *et al.*, 1982.
[c] Reactivity as described by Jackson and Webster, 1983.
NT = Not tested.

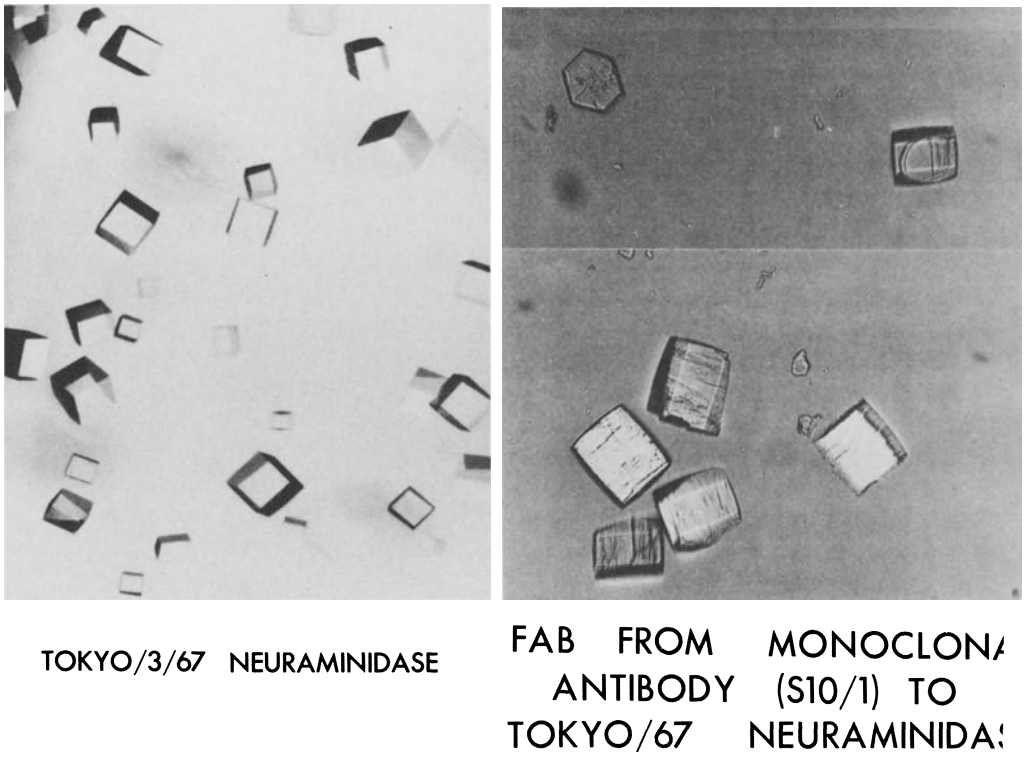


Fig. 10. Crystals, suitable for x-ray diffraction analysis, of neuraminidase heads from Tokyo/67 virus and of the Fab fragment of a monoclonal antibody (S10/1) which binds to the NA. Variant NA molecules from a virus selected in the presence of S10/1 monoclonal antibody, with a sequence change at position 368 of lysine to glutamic acid, have also been crystallized. The variant NA does not bind the S10/1 antibody and x-ray diffraction analysis will show how the sequence change affects the shape of the antigenic site. Photo of Fab crystals supplied by Peter Colman

One of the variants selected with this monoclonal antibody (S10/1) showed a sequence change at position 368 of lysine to glutamic acid (the same change occurred in the field strains in 1972). Neuraminidase heads from this variant form crystals isomorphous with those of the wild-type neuraminidase (A/Tokyo/3/67) and X-ray diffraction data from the variant crystals have been collected and are being analyzed (P. Colman, personal communication). The separate crystallization of antigen and Fab fragment opens the way to map, for the first time, the complementary surfaces of an antigen-antibody complex.

C. Antigenic Drift in the Neuraminidases of Influenza A Viruses from Lower Animals

Relatively little information is available about antigenic drift in the neuraminidases of non-human influenza viruses. Antigenic drift has been reported in swine

influenza viruses (Kendal *et al.*, 1977). Analysis of avian influenza viruses that possess N2 neuraminidase with a panel of human N2 monoclonal antibodies indicated that the avian strains were all closely related to the A/Japan/305/57 human strain (Webster *et al.*, 1982). Some antigenic variation was detectable in avian N2 strains isolated from 1965 to 1981, but there was no gradual antigenic drift similar to that found in human strains. The results indicated that multiple different strains of the same subtype co-exist and persist in the avian population. Sequences of the first 200—300 nucleotides from the 5′ ends of cDNA transcribed from NA genes of 20 avian viruses of eight of the nine NA subtypes also did not show marked drift within subtypes with time (Blok and Air, 1982 c). The situation is similar to that with HA in avian strains in that co-circulation of antigenically distinct variants of a given subtype makes detection of antigenic drift very difficult.

D. Mechanism of Antigenic Drift

Antigenic drift clearly occurs by accumulation of a series of point mutations, but since the single amino acid changes found in the antigenic variants selected with monoclonal antibodies have in most cases little effect on the antigenic properties of the HA and NA when tested with polyclonal antisera, we can ask how do variants with epidemiological potential arise in nature? Antigenic analysis of many H3N2 field isolates from 1969—1972 and from 1979—1982 with monoclonal antibodies to the HA indicate that the viruses circulating in nature are heterogeneous (unpublished results). Many different co-circulating variants can be detected during an epidemic year with monoclonal antibody to the HA (Webster *et al.*, 1979). Several of the field isolates in 1979—1982 reacted with a panel of monoclonal antibodies to the HA of A/Bangkok/1/79 in a fashion identical to laboratory-selected monoclonal variants. The laboratory-selected variants and the field isolates could not be distinguished from A/Bangkok/1/79 with post-infection ferret antisera.

It is apparent that field isolates with minor antigenic changes do circulate and cause disease and that a change is not required in each of the four antigenic regions of H3N2 viruses (Wiley *et al.*, 1981) to give an epidemiologically important virus. It appears that a change in one antigenic region may be enough for the virus to escape neutralization and cause disease.

VII. Antigenic Shift

Since the first human influenza virus was isolated in 1933, antigenic shifts in type A influenza viruses have occurred in 1957 when the H2N2 subtype (Asian influenza) replaced the H1N1 subtype, in 1968 when the Hong Kong (H3N2) virus appeared, and in 1977 when the H1N1 virus reappeared. All these major antigenic shifts in the virus occurred in China and anecdotal records suggest that previous epidemics also had their origin in China. Antigenic shifts have not been detected in influenza B viruses.

A. Evidence from Sequence Data

Although we are still uncertain how these new subtypes appear or reappear in the human population, some clues are emerging from the plethora of recent sequence data.

The complete amino acid sequences of HA1 and HA2 have been deduced from the sequences of cloned DNA copies of the genes for strains representative of four different HA subtypes of type A influenza—fowl plague (H7N7) virus (Porter *et al.*,

Fig. 11. Dendrogram showing the relationships between the HA1 N-terminal amino acid sequences deduced from genomic sequences of viruses representing the 13 known subtypes. The sequences were aligned using amino acids that are invariable in certain positions of all sequences (amino acid position from N terminus of most subtypes: CYS 4, GLY 6, THR 18, VAL 26, CYS 42, CYS 55, GLY 63, PRO 65, CYS 67, GLU 81). The percentage sequence differences of all pairwise comparisons of the aligned sequences starting from the N-terminal ASP or corresponding amino acid were used to calculate the dendrogram. Thus, the positions of each bifurcation in the dendrogram indicate the mean sequence difference of the sequences connected through that point

1979), Asian (H2N2) influenza virus (Gething *et al.*, 1980), PR8 and WSN (H1N1) viruses (Winter *et al.*, 1981; Hiti *et al.*, 1981), and a number of strains within the human Hong Kong (H3N2) subtype (Both and Sleigh, 1981; Verhoeyen *et al.*, 1980). The sequences of two strains of A/Duck/Ukraine/63 (H3N8) hemagglutinin have also been determined, one from gene sequence data (Fang *et al.*, 1981), the other from direct protein sequencing (Ward and Dopheide, 1982). A full protein sequence analysis is available only for the H3 strain (Ward and Dopheide, 1980; Dopheide and Ward, 1980; Ward, 1981), although sufficient protein data are available for the H2 strain to indicate where the signal peptide is removed and where the molecule is cleaved into HA1 and HA2. The data also show the location of the single disulphide bond connecting HA1 and HA2 and that most of the potential glycosylation sites in the HA do have carbohydrate attached.

Sequences up to 350 nucleotides from the 3′ end of the HA gene and the predicted amino acid sequences from 32 type A influenza viruses, including representatives of each of the 13 known HA subtypes, have been determined (Air, 1981; Hinshaw *et al.*, 1982). Cysteine residues and certain other amino acids are conserved in all sequences, indicating that the 13 HA subtypes evolved from a common ancestor and share a common basic structure. When the partial amino acid sequences are compared pairwise, the most distant subtypes are H1 and H3 (25% homology). Other subtypes are more similar to each other, the highest homology (80%) being between H2 and H5. Relationships between the partial sequences can be illustrated as a dendrogram (Fig. 11).

Within strains of a single subtype, the differences in deduced amino acid sequence are generally less than 10% (Air, 1981), although differences in the N-terminal region of HA1 from two strains each of H7 and H10 viruses are close to 20%, almost equal to the difference between the amino acid sequences of the two closest subtypes (H2 and H5).

Sequence data from different subtypes of NA is also accumulating. From cDNA copies of the gene cloned in E. coli, complete sequences are available for the N1 strains PR/8/34 (Fields *et al.*, 1981) and WSN/33 (Hiti and Nayak, 1982) and N2 strains A/Udorn/72 (Markoff and Lai, 1982), NT/60/68 (Bentley and Brownlee, 1982) and RI/5^{+}/57 (Ward, personal communication).

For the other NA subtypes, data is available only for the first 200 to 300 nucleotides from the 5′ end of the cDNA (Blok and Air, 1982 a). Predicted amino acid sequences show that the N-terminal 6 amino acids of NA are conserved in at least 8 subtypes, the next 6 are conserved in most subtypes, then the membrane-insertion sequence and stalk region show amazing variation, and no homology can be seen between subtypes. In the "head" region of N1 and N2 strains, considerable homology is found, suggesting that the overal structure is maintained as in the case of the HA, but data from the "head" regions of other NA subtypes are required before relationships among NA subtypes can be seen.

B. Possible Mechanisms of Shifts in Human Strains

Whereas it is possible to imagine that the H2 viruses could eventually mutate into H5, the rapid change to H3 must involve a more drastic mechanism. How, then,

could the H3 hemagglutinin have replaced the H2 hemagglutinin in the human population in 1968? This "new" human virus may be derived from animal or avian viruses by genetic reassortment. There is ample evidence for genetic reassortment between influenza A viruses from humans and lower animals *in vivo* (Webster *et al.*, 1971) and recent studies by Young and Palese, 1979; Bean *et al.*, 1980 provide evidence for reassortment among human strains. The Hong Kong H3N2 strain of human influenza has been shown to be a reassortant (Laver and Webster, 1973; Scholtissek *et al.*, 1978). This virus contains the NA (and all other) genes from an Asian (H2N2) strain of human influenza and a HA which is antigenically related to that of A/Duck/Ukraine/63 (H3N8) and A/Equine/2/Miami/63 (H3N8) viruses (Fang *et al.*, 1981; Ward and Dopheide, 1981). The amino acid sequence homology between the HAs of A/Duck/Ukraine/63 and A/Aichi/2/68 viruses (both of subtype H3) is 96%. We do not know if it was an animal or bird virus which donated the HA gene during the recombination event that led to the formation of the Hong Kong strain. The donating virus could have been one which contained Hong Kong HA that had persisted unchanged since the 1900 epidemic. It is possible that viruses such as equine (H3N8) and Duck/Ukraine (H3N8) have evolved from the 1900 epidemic strain.

Second, the "new" virus may have caused an epidemic in man many years previously and may have remained hidden and unchanged in some unknown place ever since. Evidence for this kind of event has been obtained. The strain of "Russian flu" (H1N1) which reappeared in Anshan in northern China in May 1977 and subsequently spread to the rest of the world, seems to be identical, in all genes, to the virus which caused an influenza epidemic in 1950 (Nakajima *et al.*, 1978; Scholtissek *et al.*, 1978). Where was this virus for 27 years?

The possible explanations include preservation in a frozen state, preservation in an animal reservoir, or retention in an integrated, as yet undetected, form in the genetic material of a human or lower animal. The first of these options has been tacitly accepted by some people mainly because there is no convincing alternative. The animal reservoir option is a possibility for H3N2 viruses, since they have been found in pigs many years after they disappeared from man (Shortridge *et al.*, 1977). The difficulty with this option is that the RNAs of swine and avian influenza viruses are very heterogeneous (Palese and Ritchey, 1977 a; Hinshaw *et al.*, 1978), making it unlikely that a human strain could be conserved in all genes for many years. The absence of antigenic drift over a number of years would also be difficult to explain. There is no evidence for integration of influenza genetic material into the host genome, although, under exceptional circumstances of co-infection with a virus possessing reverse transcriptase, this might be possible.

The third way in which new viruses could appear in the human population would be if an animal or bird virus became infectious for man. This could be brought about by mutation, but the number of mutations required and the genes in which these would need to occur are unknown. Periodical transmission of swine influenza viruses to humans evidently occurs. Witness the incident in soldiers at Fort Dix and the isolation of virtually identical viruses from pigs and man on the same farm in Wisconsin (Hinshaw *et al.*, 1978). Most of these transmissions are "dead end" in that the viruses have little or no capacity to transmit to human

contacts and initiate an epidemic. However, the disastrous pandemic virus of 1918–1919 did possess this property.

Perhaps each of these mechanisms has operated at one time or another to produce a "major shift" in influenza viruses infecting man. We do not know how the different subtypes of subtype A influenza arose, but it is reasonable to suppose that "shuffling" of genes between influenza A strains is occurring all the time, most frequently in birds (Hinshaw *et al.*, 1980), but also in mammals, including man (Young and Palese, 1979; Bean *et al.*, 1980).

C. Antigenic Shift in Influenza Viruses from Lower Animals

Seals: Although the appearance of influenza virus in seals probably resulted from host range variation, the event was of considerable biological importance for it has implications on how new pandemic strains may appear. In 1979–1980, approximately 20% of the harbour seal *(Phoca uitulina)* population of the northeast coast of the United States died of a severe respiratory infection with consolidation of the lungs, typical of primary viral pneumonia (Geraci *et al.*, 1982). Influenza virus particles were found in high concentrations in the lungs and brains of the dead seals. Antigenic analysis showed that this virus was closely related to fowl plague virus (A/FPV/Dutch/27) (H7N7), a highly lethal influenza virus of chickens not previously found in mammals (Webster *et al.*, 1981). Analysis of the RNAs of the seal virus by competitive RNA-RNA hybridization showed that all of the genes were closely related to those from different avian influenza strains. It is not known whether this virus originated by transmission from birds or whether influenza in seals has previously escaped detection. Serological and biological information favors the former explanation—no serological evidence for seal influenza was obtained except in the surviving animals on the New England coast, and since 1980 there has been no further evidence of this virus in seals (Geraci, personal communication).

The A/Seal/Mass/1/80 (H7N7) influenza virus provides the first evidence that a strain deriving all of its genes from one or more avian influenza viruses can be associated with severe disease in a mammalian population in nature. Whether this breach of species specificity represents a unique event in influenza evolution remains to be determined, but the occurrence raises the possibility that human or animal influenza viruses may be derived directly from avian strains. If this event had occurred in humans instead of seals, the resulting pandemic might have been similar to that in 1918–1919.

Other animals: Since representatives of each of the known subtypes of influenza A viruses cocirculate in avian species (Table 1) antigenic shift cannot be detected. There is no doubt that genetic reassortment between influenza A viruses is a common occurrence in avian strains (Desselberger *et al.*, 1978; Hinshaw *et al.*, 1980) and almost all isolates of a given subtype are genetically diverse (Sriram *et al.*, 1980).

Although two different subtypes of influenza viruses cocirculate in horses (Table 1) there is no evidence for genetic reassortment between these viruses. In

pigs, where swine influenza viruses (H1N1) persist and H3N2 viruses either persist or are periodically introduced from humans, reassortants possessing H1N2 (Hsw1N2) have recently been detected in Japan (Sugimura *et al.*, 1980). The isolation of the virus demonstrates that antigenic shift can occur in influenza virus in pigs but to date there is no evidence that this virus has become an epidemic strain.

VIII. Variation in the Nucleoprotein

The nucleoprotein (NP) is one of the group-specific antigens of influenza viruses that distinguishes between the influenza A, B, and C viruses. It probably constitutes the backbone of the helical internal complex that is associated with the RNA segments and the three different polymerases.

The NP gene of A/PR/8/34 virus is 1,565 nucleotides long and is capable of encoding a protein of 498 amino acid residues (MW 56,106) which is rich in arginine (Winter and Fields, 1981). Double immunodiffusion tests showed antigenic differences between the NPs of H1N1 and the H3N2 strains (Schild *et al.*, 1979) and studies with monoclonal antibodies to the NP of A/WSN/33 (H1N1) viruses have shown that antigenic variation occurs in this molecule (van Wyke *et al.*, 1980). The NP molecule possesses at least three nonoverlapping antigenic areas, one area being the same on all strains tested. Monoclonal antibodies to this conserved domain inhibited *in vitro* transcription of viral RNA, suggesting that this region of the NP is involved in RNA transcription (van Wyke *et al.*, 1981).

Genetic analysis of a large number of influenza strains by competitive RNA-RNA hybridization has revealed that the nucleoprotein genes of each of these viruses can be placed into one of five different nucleoprotein groupings (Bean, personal communication). All of the avian strains fall into two groups, the equine virus strains form two other groups and one group contains all of the human and swine influenza viruses. This restriction of groups of genetically related nucleoprotein genes to viruses infecting certain species suggests that this protein may play a role in determining species specificity.

IX. Variation in the Matrix Protein

The complete sequence of RNA segment 7 (M) of two strains of A/PR/8/34 (H1N1) (Winter and Fields, 1980; Allen *et al.*, 1980) and of A/Udorn/72 (H3N2) (Lamb and Lai, 1981) has been reported as well as partial sequences of a number of strains (Hall and Air, 1981). Following the first AUG codon in the positive strand, a 252-residue protein, rather hydrophobic and rich in arginine, is encoded. Compositions of tryptic peptides from purified matrix protein of PR8 (Hall and Air, 1981) correspond well with those expected from the amino acid sequence predicted from the gene sequence.

Comparison of the sequences of RNA segment 7 of the H3N2 (Udorn) and H1N1 (PR8) strains shows that the sequences coding for the matrix (M) proteins of these viruses isolated 38 years apart are highly conserved (Lamb and Lai, 1981), in

keeping with antigenic studies (Schild, 1972). Comparison of 230 nucleotides of RNA segment 7 from five human H1N1, H2N2 and H3N2 strains isolated over a 43-year period suggests that the same segment 7 was retained throughout the antigenic shifts of HA and NA type (H1N1 to H2N2 to H3N2) (Hall and Air, 1981). In addition, the complete sequences contain a second open reading frame which overlaps the M protein sequence by 68 nucleotides.

Three mRNAs transcribed from RNA segment 7 have been isolated. One (M1 mRNA) consists of an uninterrupted, nearly full-length copy of RNA segment 7 and is responsible for the production of the M protein. An M2 protein is generated from a spliced product of the M1 mRNA, such that, following the nucleotides encoding the N-terminal nine amino acids, nearly 600 nucleotides are spliced out and the reading frame is changed; a protein product corresponding to this reading frame has been identified in infected cells (Lamb and Choppin, 1981). In addition, a third mRNA has been found, which would code only for an 8-residue peptide identical to the C terminus of M1 (Lamb *et al.*, 1981). Such a product has not yet been isolated.

Studies with a panel of monoclonal antibodies to the matrix protein (M1) of A/WSN/33 (H1N1) indicates that there are 4 antigenic sites on this molecule (van Wyke, personal communication) and that antigenic differences can be detected both within and between subtypes.

X. Variation in the Nonstructural Proteins

Recent studies have shown that RNA segment 8 codes for at least two nonstructural polypeptides, NS1 and NS2, which are translated from separate mRNAs (Lamb and Lai, 1980; Porter *et al.*, 1980; Inglis *et al.*, 1979). Mapping and sequence studies have shown that NS1 and NS2 overlap by 70 amino acids that are translated from different reading frames. Polypeptides NS1 and NS2 share 9 amino acids at their N termini, but after this sequence the mRNA for NS2 has a deletion of 423 nucleotides, then rejoins the rest of the mRNA in the +1 reading frame. The function of NS1 or NS2 has not been established. NS1 is made in large amounts and accumulates in the nucleus (Lazarowitz *et al.*, 1971); NS2 is made late in infection and is found predominantly in the cytoplasm (Lamb and Lai, 1980; Mahy *et al.*, 1980).

Because NS1 and NS2 are internal proteins of infected cells and hence less available to antibodies, they would be expected to show less sequence variation than the surface glycoproteins (HA and NA). Accordingly, comparison of the sequences of the NS genes of fowl plague (Porter *et al.*, 1980) and the two human influenza strains A/Udorn/72 (H3N2) and A/PR/8/34 shows only 8—11% differences (Lamb and Lai, 1980; Baez *et al.*, 1980; Winter *et al.*, 1981). An open reading frame potentially coding for a polypeptide has been noted in the noncoding, virion RNA of the NS genes of A/PR/8/34, Udorn/72 and FPV, but is not present in the NS gene of Duck/Alberta/60/76 (Baez *et al.*, 1981). No protein corresponding to this extra "gene" has yet been identified.

The NS1 protein has been found to be very heterogeneous with respect to charge and phosphorylation (Petri *et al.*, 1982). The degree of phosphorylation appears to be characteristic of particular subtypes. Antigenic analysis of NS1 products from different subtypes with polyclonal antisera indicated significant cross reactivity between all influenza A viruses but a gradual antigenic drift was detectable in competitive RIA assays (Shaw *et al.*, 1982).

XI. Variation in the Polymerase Proteins

The three largest proteins of the virion (PB1, PB2, PA) with MWs of 96,000, 87,000, and 85,000, respectively, are found in association with the nucleoprotein and virion RNA and carry the polymerase activity (Bishop *et al.*, 1972) which transcribes the invading viral RNA (Bean and Simpson, 1973). Proteins PB1 and PB2 are probably required for complementary RNA synthesis and PA and NP for virion RNA synthesis (Palese and Ritchey, 1977 b). To date no information is available about antigenic variation in the P proteins.

Transcription of the viral RNA segments soon after infection terminates ~40 nucleotides before the 5′ end to give messenger RNAs (Hay and Skehel, 1982); these mRNAs have a cap structure and a heterogeneous sequence of 10 to 13 nucleotides derived from cellular mRNAs at their 5′ end (Plotch *et al.*, 1979). Later transcripts are full length and act as templates for synthesis of the genes of the progeny virus. The complete nucleotide sequence of the three polymerase genes of the A/PR/8/34 strain (Fields and Winter, 1982), and at least two to the P genes of A/NT/60/68 (Bishop *et al.*, 1982) have been determined, but the extent of variation in the polymerase genes is unknown, although these may play an important part in host range and virulence.

Defective interfering (DI) influenza virus particles are generated by high multiplicity passage in permissive cells. These particles are of interest because they facilitate the establishment of persistent infection in cell cultures and could therefore engender latency. They contain small RNA molecules which are absent from infectious virus and which are generated predominantly by massive internal deletion from the three P genes (Nayak, 1980; Nakjima *et al.*, 1979), although the cloned DNA of one small RNA appears to be a mosaic of several segments from at least two of the polymerase genes (Fields and Winter, 1982; Moss and Brownlee, 1981).

XII. What of the Future?

The recent explosion of structural data on the genes and gene products of influenza viruses together with information from monoclonal antibodies has provided some important information about the behavior of this virus. For example, it is clear that antigenic shift does not occur by direct mutation of one subtype into another, whereas antigenic drift occurs by point mutations in the genes. However, the mutation frequency of the surface proteins is not significantly greater than that of the

internal proteins. HA possesses at least four distinct antigenic sites which can vary independently, with pathogenicity depending in part on the sequence in the HA where cleavage occurs. Some RNA segments of the virus possess overlapping genes, using two reading frames. The sequence data support the antigenic classification of the HA and NA subtypes.

However, many mysteries remain. Shifts can occur by the emergence of viruses which previously caused epidemics and have remained unchanged, as if frozen, for many years. Where these viruses hide and what causes them to reemerge are mysteries. The shifts may also occur by genetic reassortment of human and animal influenza viruses or by mutation of an animal or avian virus so that it becomes able to infect people. Most type A influenza virus subtypes are found in birds, but the role of these in the emergence (or reemergence) of new subtypes in man is not clear; nor is it clear why the shifts seem always to occur in China or why the old subtypes usually disappear from man as the pandemic of new virus starts. On the other hand, the Hong Kong subtype did not disappear when Russian flu started and we do not know why this occurred.

There are many unanswered questions about the origin of new pandemic strains of influenza virus and the molecular basis of antigenic variation (Webster *et al.*, 1982). Why does antigenic drift occur in influenza, but not, for example, in paramyxoviruses such as measles virus, where antigenic variants can be selected *in vitro* with monoclonal antibodies at about the same frequency as influenza (Portner *et al.*, 1980; Birrer *et al.*, 1981)? Why does one sometimes observe "original antigenic sin", (Francis, 1960) in which a virus may preferentially elicit antibodies against a previous strain which has since disappeared? In addition to antigenic variation, variation in host range, virulence and transmissibility greatly influence the severity of the disease and the spread of epidemics. Is it possible that a new virus could appear which would cause the same degree of mortality as the 1918 "swine" flu epidemic? What influences the spread of influenza virus, and why do epidemics normally appear in the winter, thus oscillating between the Northern and Southern Hemispheres? It is not known why epidemics are self-limiting, even when susceptible hosts remain in abundance.

There are now, however, better prospects for the control of influenza. Effective and safe live vaccines may be made in the future using modern genetic engineering techniques. Synthetic vaccines may be possible where the antigens are pure enough for the vaccines to be given in large enough doses to be effective and engineered such that they do not induce original antigenic sin. A universal vaccine, which protects against all members of a subtype or even across subtypes might be feasible. It might also be more rewarding to look for ways of increasing the cellular immune response as well as the humoral response.

With the three-dimensional structure of at least one HA molecule known and candidates for the antigen and receptor binding sites tentatively identified, the way is open for the synthesis of vaccines containing key segments of HA molecules. HA antigenic determinants can be detected in *Escherichia coli* transformed with a plasmid carrying a cloned HA gene (Emtage *et al.*, 1980; Heiland and Gething, 1981; Davis *et al.*, 1981), and copious amounts of HA appear on the surfaces of eukaryotic cells infected with a simian virus 40 (SV40) vector containing the HA gene (Gething

and Sambrook, 1981; Sveda and Lai, 1981). These experiments open the way to the use of site-directed mutagenesis to tailor the HA into a successful synthetic and possibly universal vaccine.

Knowledge of the structure of the HA molecule has also permitted the chemical synthesis of peptides for use as vaccines (Müller *et al.*, 1982; Green *et al.*, 1982; Jackson *et al.*, 1982). Many of the peptides are antigenic when coupled with protein carriers, but to date these synthetic peptides have not been shown to induce significant protection against influenza virus challenge *in vivo*. The potential clearly exists for such vaccines, for cell-mediated studies indicate the existence of cross-reactive determinants between influenza A subtypes.

What of the immediate future? We do not know if another antigenic shift is imminent. If a new virus does suddenly appear, will the structural data tell us where it came from? The new virus might not have the capacity to kill people to the extent that the recent seal virus killed seals, but if it did, it is doubtful that we would be able to stop it from causing worldwide devastation. If no shift occurs, what will happen to the currently circulating H3N2 and H1N1 strains? Will they continue to drift, or have they reached their limits? There is little doubt that a greater understanding of the molecular biology and genetics of the virus will make an essential contribution to piecing together the answers to the puzzle of influenza.

Acknowledgements

The authors were supported in part by grants AI-08831, AI-02649, and AI-15343 from the National Institute of Allergy and Infectious Diseases, and by grant CA13148 from the National Cancer Institute. The writing of this joint review was greatly assisted by international telephone facilities donated by the Australian Overseas Telecommunications Commission.

The authors thank Drs. Virginia Hinshaw, Clayton Naeve and William Bean for their assistance with the manuscript and Glenith Miller for typing the manuscript.

References

Air, G. M.: Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus. Proc. Natl. Acad. Sci. U.S.A. *78*, 7639–7643 (1981).

Allen, H., McCauley, J., Waterfield, M., Gething, M. J.: Influenza virus RNA segment 7 has the coding capacity for two polypeptides. Virology *107*, 548–551 (1980).

Baez, M., Taussig, R., Zazra, J. J., Young, J. F., Palese, P., Reisfeld, A., Skalka, A. M.: Complete nucleotide sequence of the influenza A/PR/8/34 virus NS gene and comparison with the NS genes of the A/Udorn/72 and A/FPV/Rostock/34 strains. Nucl. Acids Res. *8*, 5845–5858 (1980).

Baez, M., Zazra, J. J., Elliott, R. M., Young, J. F., Palese, P.: Nucleotide sequence of the influenza A/Duck/Alberta/60/76 virus NS RNA: Conservation of the NS1/NS2 overlapping gene structure in a divergent influenza virus RNA segment. Virology *113*, 397–402 (1981).

Bean, W. J., Simpson, R. W.: Primary transcription of the influenza virus genome in permissive cells. Virology *56*, 646–651 (1973).

Bean, W. J., Cox, N. J., Kendal, A. P.: Recombination of human influenza A viruses in nature. Nature *284*, 638–640 (1980).

Bentley, D. R., Brownlee, G. G.: Sequence of the N2 neuraminidase from influenza virus A/NT/60/68. Nucl. Acids Res. *10*, 5033–5042 (1982).

Birrer, M. J., Udem, S., Nathenson, S., Bloom, B. R.: Antigenic variants of measles virus. Nature *293,* 67–69 (1981).

Bishop, D. H., Roy, P., Bean, W. J., Simpson, R. W.: Transcription of the influenza ribonucleic acid genome by a virion polymerase. III. Completeness of the transcription process. J. Virol. *10,* 689–697 (1972).

Bishop, D. H., Huddleston, J. A., Brownlee, G. G.: The complete sequence of RNA segment 2 of influenza A/NT/60/68 and its encoded P1 protein. Nucl. Acids Res. *10,* 1335–1343 (1982).

Blok, J.: Sequence variation at the 3′ ends of neuraminidase gene segments within and among the different NA subtypes. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., Fox, C. F., eds.) (ICN-UCLA Symposia on Molecular and Cellular Biology, Vol. XXI) (Nayak, D. P., Fox, C. F., eds.), 45–54. New York: Academic Press 1981.

Blok, J., Air, G. M.: Variation in the membrane-insertion and "stalk" sequences in eight subtypes of influenza type A virus neuraminidase. Biochemistry *21,* 4001–4007 (1982 a).

Blok, J., Air, G. M.: Block deletions in the neuraminidase genes from some influenza A viruses of the N1 subtype. Virology *118,* 229–234 (1982 b).

Blok, J., Air, G. M.: Sequence variation at the 3′end of the neuraminidase gene from 39 influenza type A viruses. Virology *121,* 211–229 (1982 c).

Both, G. W., Sleigh, M. J.: Conservation and variation in the hemagglutinins of Hong Kong subtype influenza viruses during antigenic drift. J. Virol. *39,* 663–672 (1981).

Caton, A. J., Brownlee, G. G., Yewdell, J. W., Gerhard, W.: The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell *31,* 417–427 (1982).

Colman, P. M, Laver, W. G.: The structure of influenza virus neuraminidase heads at 5 Å resolution. In: Proceedings, 7th Aharan Katzir-Katchalsky Conference on Structural Aspects of Recognition and Assembly in Biological Macromolecules (Balasan, M., ed.), 869–872. Jerusalem: International Science Services 1981.

Curry, R. L., Brown, J. D., Baker, F. A., Hobson, D.: Serological studies with purified neuraminidase antigens of influenza B viruses. J. Hyg. *72,* 197–204 (1974).

Davenport, F. M., Hennessy, A. V., Francis, T., jr.: Epidemiologic and immunologic significance of age distribution of antibody to antigenic variants of influenza virus. J. Exp. Med. *98,* 641–656 (1953).

Davenport, F. M., Hennessy, A. V., Drescher, J., Webster, R. G.: Analytical, serologic and clinical experiences with the hemagglutinating subunits of influenza A virus. In: Ciba Symposium on Cellular Aspects of Myxovirus Infection, 272–287, 1964.

Davis, A. R., Nayak, D. P., Ueda, M., Hiti, A. L., Dowbenko, D., Kleid, D. G.: Expression of antigenic determinants of the hemagglutinin gene of a human influenza virus in *Escherichia coli.* Proc. Natl. Acad. Sci. U.S.A. *78,* 5376–5380 (1981).

Desselberger, U., Nakajima, K., Alfino, P., Pedersen, F. S., Haseltine, W. A., Hannoun, C., Palese, P.: Biochemical evidence that "new" influenza virus strains in nature may arise by recombination (reassortment). Proc. Natl. Acad. Sci. U.S.A. *75,* 3341–3345 (1978).

Dopheide, T. A., Ward, C. W.: Structural studies on a Hong Kong influenza hemagglutinin. The structure of the light chain and the arrangement of the disulphide bonds. In: Structure and Variation in Influenza Virus (Laver, W. G., Air, G. M., eds.), 21–26. North-Holland/Elsevier 1980.

Emtage, J. S., Tacon, W. C., Catlin, G. H., Jenkins, B., Porter, A. G., Carey, N. H.: Influenza antigenic determinants are expressed from hemagglutinin genes cloned in *Escherichia coli.* Nature *283,* 171–174 (1980).

Fang, R., Min Jou, W., Huylebroeck, D., Devos, R., Fiers, W.: Complete structure of A/Duck/Ukraine/63 influenza hemagglutinin gene: Animal virus as progenitor of human H3 Hong Kong 1968 influenza hemagglutinin. Cell *25,* 315–323 (1981).

Fields, S., Winter, G., Brownlee, G. G.: Structure of the neuraminidase gene from human influenza virus A/PR/8/34. Nature (Lond.) *290,* 213–217 (1981).

Fields, S., Winter, G.: Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structure of a small viral RNA segment. Cell *28,* 303–313 (1982).

Francis, T.: On the doctrine of original antigenic sin. Proc. Am. Phil. Soc. *104,* 572 (1960).

Geraci, J. R., St. Aubin, D. J., Barker, I. K., Webster, R. G., Hinshaw, V. S., Bean, W. J., Ruhnke, H. L., Prescott, J. H., Early, G., Baker, A. S., Madoff, S., Schooley, R. T.: Mass mortality of harbor seals: Pneumonia associated with influenza A virus. Science *215,* 1129–1131 (1982).

Gerhard, W., Yewdell, J., Frankel, M., Webster, R. G.: Antigenic structure of influenza virus hemagglutinin defined by hybridoma antibodies. Nature *290,* 713–717 (1981).

Gething, M. J., White, J. M., Waterfield, M. D.: Purification of the fusion protein of Sendai virus: Analysis of the NH_2-terminal sequence generated during precursor activation. Proc. Natl. Acad. Sci. U.S.A. *75,* 2737–2740 (1978).

Gething, M. J., Bye, J., Skehel, J., Wakefield, M.: Cloning and DNA sequence of double-stranded copies of hemagglutinin genes from H2 and H3 strains elucidates antigenic shift and drift in human influenza virus. Nature *287,* 301–306 (1980).

Gething, M. J., Sambrook, J.: Cell-surface expression of influenza hemagglutinin from a cloned DNA copy of the RNA gene. Nature *293,* 620–625 (1981).

Green, N., Alexander, H., Olson, A., Alexander, S., Shinnick, T. M., Sutcliffe, J. G., Lerner, R. A.: Immunogenic structure of the influenza virus hemagglutinin. Cell *28,* 477–487 (1982).

Hall, R. M., Air, G. M.: Variation in nucleotide sequences coding for the N-terminal regions of the matrix and nonstructural proteins of influenza A viruses. J. Virol. *38,* 1–7 (1981).

Hay, A. J., Skehel, J. J.: The structure and replication of influenza viruses. In: Basic and Applied Influenza Research (Beare, A. S., ed.), 105–118. CRC Press 1982.

Heiland, I., Gething M. J.: Cloned copy of the hemagglutinin gene codes for human influenza antigenic determinants in *E. coli.* Nature *292,* 851–852 (1981).

Hinshaw, V. S., Bean, W. J., jr., Webster, R. G., Easterday, B. C.: The prevalence of influenza viruses in swine and the antigenic and genetic relatedness of influenza viruses from man and swine. Virology *84,* 51–62 (1978).

Hinshaw, V. S., Bean, W. J., Webster, R. G., Sriram, G.: Genetic reassortment of influenza A viruses in the intestinal tract of ducks. Virology *102,* 412–419 (1980).

Hinshaw, V. S., Webster, R. G., Turner, B.: The perpetuation of orthomyxoviruses and paramyxoviruses in Canadian waterfowl. Can. J. Micro. *26,* 622–629 (1980).

Hinshaw, V. S., Webster, R. G., Bean, W. J., Sriram, G.: The ecology of influenza viruses in ducks and analysis of influenza viruses with monoclonal antibodies. Comp. Immunol. Microbiol. Infect. Dis. *3,* 155–164 (1981).

Hinshaw, V. S., Air, G. M., Gibbs, A. J., Graves, L., Prescott, B., Karunakaran, D.: Antigenic and genetic characterization of a novel hemagglutinin subtype of influenza A viruses from gulls. J. Virol. *42,* 865–872 (1982).

Hinshaw, V. S., Naeve, C. W., Webster, R. G., Douglas, A., Skehel, J. J., Bryans, J.: Analysis of antigenic variation in equine 2 influenza viruses. Bull. Wld. Hlth. Org. *61,* 153–158 (1983).

Hirsch, A.: Handbook of Geographical and Historical Pathology, Vol. 1, 7. London: New Sydenham Society 1883.

Hiti, A. L., Davis, A. R., Nayak, D. P.: Complete sequence analysis shows that the hemagglutinins of the H0 and H2 subtypes of human influenza virus are closely related. Virology *111,* 113–124 (1981).

Hiti, A. L., Nayak, D. P.: Complete nucleotide sequence of the neuraminidase gene of human influenza virus A/WSN/33. J. Virol. *41,* 730–734 (1982).

Huang, R. T. C., Rott, R., Klenk, H. D.: Influenza viruses cause hemolysis and fusion of cells. Virology *110,* 243–247 (1981).

Inglis, S. C., Barrett, T., Brown, C. M., Almond, J. W.: The smallest genome RNA segment of influenza virus contains two genes that may overlap. Proc. Natl. Acad. Sci. U.S.A. *76,* 3790–3794 (1979).

Jackson, D. C., Murray, J. M., White, D. O., Fagan, C. N., Tregear, G. W.: Antigenic acitivity of a synthetic peptide comprising the "loop" region of influenza virus hemagglutinin. Virology *120,* 273–276 (1982).

Jackson, D. C., Webster, R. G.: A topographical map of the enzyme active center and antigenic sites on the neuraminidase of influenza virus A/Tokyo/3/67 (H2N2). Virology *123,* 69–77 (1982).

Kendal, A. P., Kiley, M. P.: Characterization of influenza virus neuraminidases: Peptide changes associated with antigenic divergence between early and late N2 neuraminidases. J. Virol. *12,* 1482–1490 (1973).

Kendal, A. P., Noble, G. R., Dowdle, W. R.: Swine influenza viruses isolated in 1976 from man and pig contain two coexisting subpopulations with antigenically distinguishable hemagglutinins. Virology *82,* 111–121 (1977).

Kendal, A., Cox, N., Nakajima, S., Webster, R. G., Bean, W. J., jr., Beare, P.: Natural and unnatural variation in influenza A (H1N1) viruses since 1977. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., Fox, C. F., eds.) (ICN-UCLA Symposia on Molecular and Cellular Biology, Vol. XXI), 489–504. New York: Academic Press 1981.

Kida, H., Yanagawa, R., Matsuoka, Y.: Duck influenza lacking evidence of disease signs and immune response. Infect. Immun. *30,* 547–553 (1980).

Lamb, R., Lai, C. J.: Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. Cell *21,* 475–485 (1980).

Lamb, R., Choppin, P. W.: Identification of a second protein (M_2) encoded by RNA segment 7 of influenza virus. Virology *112,* 729–737 (1981).

Lamb, R. A., Lai, C. J.: 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 (1981).

Lamb, R. A., Lai, C. J., Choppin, P. W.: Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: Colinear and interrupted mRNAs code for overlapping proteins. Proc. Natl. Acad. Sci. U.S.A. *78,* 4170–4174 (1981).

Laver, W. G., Valentine, R. C.: Morphology of the isolated hemagglutinin and neuraminidase subunits of influenza virus. Virology *38,* 105–119 (1969).

Laver, W. G., Webster, R. G.: Studies on the origin of pandemic influenza. III. Evidence implicating duck and equine influenza viruses as possible progenitors of the Hong Kong strain of human influenza. Virology *51,* 383–391 (1973).

Laver, W. G.: Crystallization and peptide maps of neuraminidase "heads" from H2N2 and H3N2 influenza virus strains. Virology *86,* 78–87 (1978).

Laver, W. G., Air, G. M., Webster, R. G., Gerhard, W., Ward, C. W., Dopheide, T. A. A.: Antigenic drift in type A influenza virus: Sequence differences in the hemagglutinin of Hong Kong (H3N2) variants selected with monoclonal hybridoma antibodies. Virology *98,* 226–237 (1979).

Laver, W. G., Air, G. M., Webster, R. G.: Mechanism of antigenic drift in influenza virus. Amino acid sequence changes in an antigenically active region of Hong Kong (H3N2) influenza virus hemagglutinin. J. Mol. Biol. *145,* 339–361 (1981).

Laver, W. G., Air, G. M., Webster, R. G., Markoff, L. J.: Amino acid sequence changes in antigenic variants of type A influenza virus N2 neuraminidase. Virology *122,* 450–460 (1982).

Lazarowitz, S. G., Compans, R. W., Choppin, P. W.: Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes. Virology *46,* 830–843 (1971).

Maeda, T., Ohnishi, S.: Activation of influenza virus by acidic media causes hemolysis and fusion of erythrocytes. FEBS Lett. *122,* 283–287 (1980).

Mahy, B. W. J., Barrett, T., Briedis, D. J., Brownson, J. M., Wolstenholme, A. J.: Influence of the host cell on influenza virus replication. Phil. Trans. R. Soc. *288,* 349–357 (1980).

Markoff, L., Lai, C. J.: Sequence of the influenza A/Udorn/72 (H3N2) virus neuraminidase gene as determined from cloned full-length DNA. Virology *119,* 288–297 (1982).

Masurel, N.: Antibody response obtained by vaccination with the influenza A/Equi 2 virus in man. Nature (Lond.) *218,* 100–101 (1968).

Moss, B. A., Brownlee, G.: Sequence of DNA complementary to a small RNA segment of influenza virus A/NT/60/68. Nucl. Acids Res. *9,* 1941–1947 (1981).

Mulder, J., Masurel, N.: Pre-epidemic antibody against 1957 strain of Asiatic influenza in serum of older people living in the Netherlands. Lancet *I,* 810–814 (1958).

Müller, G. M., Shapira, M., Arnon, R.: Anti-influenza response achieved by immunization with a synthetic conjugate. Proc. Natl. Acad. Sci. U.S.A. *79,* 569–573 (1982).

Nakajima, K., Desselberger, U., Palese, P.: Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. Nature *274,* 334–339 (1978).

Nakajima, K., Ueda, M., Sugiura, A.: Origin of small RNA in von Magnus particles of influenza virus. J. Virol. *29,* 1142–1148 (1979).

Nayak, D. P.: Defective interfering influenza viruses. Ann. Rev. Microbiol. *34,* 619–644 (1980).

Noble, G. R.: Epidemiological and clinical aspects of influenza. In: Basic and Applied Influenza Research (Beare, A. S., ed.), 11–50. CRC Press 1982.

Palese, P., Ritchey, M. B.: Polyacrylamide gel electrophoresis of the RNAs of new influenza virus

strains: An epidemiological tool. In: Proceedings of International Symposium on Influenza Immunization. II. Develop. Biol. Standard *39,* 411–414 (1977 a).

Palese, P., Ritchey, M. B.: Live attenuated influenza virus vaccines. Strains with temperature-sensitive defects in P3 protein and nucleoprotein. Virology *78,* 183–191 (1977 b).

Paniker, C. K.: Serological relationships between the neuraminidase in influenza viruses. J. Gen. Virol. *2,* 385–394 (1968).

Pereira, H. G.: Influenza: Antigenic spectrum. Progr. Med. Virol. *11,* 46–79 (1969).

Pereira, M. S.: Global surveillance of influenza. Brit. Med. Bull. *35,* 9–14 (1979).

Petri, T., Patterson, S., Dimmock, N. J.: Polymorphism of the NS1 proteins of type A influenza virus. J. Gen. Virol. *61,* 217–231 (1982).

Plotch, S. J., Bouloy, M., Krug, R. M.: Transfer of 5′-terminal cap of globin mRNA to influenza viral complementary RNA during transcription *in vitro.* Proc. Natl. Acad. Sci. U.S.A. *76,* 1618–1622 (1979).

Porter, A. G., Barber, C., Carey, N. H., Hallewell, R. A., Threlfall, G., Emtage, J. S.: Complete nucleotide sequence of an influenza virus hemagglutinin gene from cloned DNA. Nature *282,* 471–477 (1979).

Porter, A. G., Smith, J. C., Emtage, J. S.: Nucleotide sequence of influenza virus RNA segment 8 indicates that coding regions for NS_1 and NS_2 proteins overlap. Proc. Natl. Acad. Sci. U.S.A. *77,* 5074– 5078 (1980).

Portner, A., Webster, R. G., Bean, W. J.: Similar frequencies of antigenic variants in Sendai, vesicular stomatitis and influenza A viruses. Virology *104,* 235–238 (1980).

Russ, G., Gerhard, W., Laver, W. G.: The antigenic topology of the hemagglutinin molecule of influenza virus A/PR/8/34. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., Fox, C. F., eds.) (ICN-UCLA Symposia on Molecular and Cellular Biology, Vol. XXI), 297–307. New York: Academic Press 1981.

Scheid, A., Choppin, P. W.: Two disulfide-linked polypeptide chains constitute the active F protein of paramyxoviruses. Virology *80,* 54–66 (1977).

Schild, G. C., Newman, R. W.: Immunological relationships between the neuraminidase of human and animal influenza viruses. Bull. Wld. Hlth. Org. *41,* 437–445 (1969).

Schild, G. C.: Evidence for a new type-specific structural antigen of the influenza virus particle. J. Gen. Virol. *15,* 99–103 (1972).

Schild, G. C., Oxford, J. S., Newman, R. W.: Evidence for antigenic variation in influenza A nucleoprotein. Virology *93,* 569–573 (1979).

Scholtissek, C., Rohde, W., Harms, E., Rott, R.: Assignment of gene functions to RNA segments of influenza A2-Singapore and genetic relatedness to other influenza strains. In: Negative Strand Viruses and the Host Cell (Barry, R. D., Mahy, B. W. J., eds.), 19–25. London: Academic Press 1978.

Scholtissek, C., van Hoyningen, V., Rott, R.: Genetic relatedness between the new 1977 epidemic strains (H1N1) of influenza and human influenza strains isolated between 1947 and 1957. Virology *89,* 613–617 (1978).

Shaw, M. W., Lamon, E. W., Compans, R. W.: Immunologic studies of the influenza A virus nonstructural protein NS1. J. Exp. Med. *156,* 243–254 (1982).

Shortridge, K. F., Webster, R. G., Butterfield, W. K., Campbell, C. H.: Persistence of Hong Kong influenza virus variants in pigs. Science *196,* 1454–1455 (1977).

Skehel, J. J., Waterfield, M. D.: Studies on the primary structure of the influenza virus hemagglutinin. Proc. Natl. Acad. Sci. U.S.A. *72,* 93–97 (1975).

Skehel, J. J., Bayley, P. M., Brown, E. B., Martin, S. R., Waterfield, M. D., White, J. M., Wilson, I. A., Wiley, D. C.: Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. Proc. Natl. Acad. Sci. U.S.A. *79,* 968–972 (1982).

Sriram, G., Bean, W. J., jr., Hinshaw, V. S., Webster, R. G.: Genetic diversity among avian influenza viruses. Virology *105,* 592–599 (1980).

Sugimura, T., Yonemochi, H., Ogawa, T., Tanaka, Y., Kumagai, T.: Isolation of a recombinant influenza virus (HswlN2) from swine in Japan. Arch. Virol. *66,* 271–274 (1980).

Sveda, M. M., Lai, C. J.: Functional expression in primate cells of cloned DNA coding for the hemagglutinin surface glycoprotein of influenza virus. Proc. Natl. Acad. Sci. U.S.A. *78,* 5488–5492 (1981).

Van Wyke, K. L., Hinshaw, V. S., Bean, W. J., jr., Webster, R. G.: Antigenic variation of influenza A virus nucleoprotein detected with monoclonal antibodies. J. Virol. *35,* 24–30 (1980).

Van Wyke, K. L., Bean, W. J., jr., Webster, R. G.: Monoclonal antibodies to the influenza A virus nucleoprotein affecting RNA transcription. J. Virol. *39,* 313–317 (1981).

Verhoeyen, M., Fang, R., Min Jou, W., Devos, R., Huylebroeck, D., Saman, E., Fiers, W.: Antigenic drift between the hemagglutinin of the Hong Kong influenza strains A/Aichi/2/68 and A/Victoria/3/75. Nature *286,* 771–776 (1980).

Ward, C. W., Dopheide, T. A.: The Hong Kong (H3) hemagglutinin. Complete amino acid sequence and oligosaccharide distribution for the heavy chain of A/Memphis/102/72. In: Structure and Variation in Influenza Virus (Laver, W. G., Air, G. M., eds.), 27–38. North-Holland/Elsevier 1980.

Ward, C. W.: Structure of the influenza virus hemagglutinin. Curr. Topics Microbiol. Immunol. *94/95,* 1–74 (1981).

Ward, C. W., Dopheide, T. A.: Amino acid sequence and oligosaccharide distribution of the hemagglutinin from an early Hong Kong influenza virus variant A/Aichi/2/68 (X-31). Biochem. J. *193,* 953–962 (1981).

Ward, C., Dopheide, T.: The Hong Kong hemagglutinin. Structural relationships between the human (H3) hemagglutinin and the hemagglutinin from the putative progenitor strain A/Duck/Ukraine/1/63 (Hav7). In: Genetic Variation Among Influenza Viruses (Nayak, D. P., Fox, C. F., eds.), (ICN-UCLA Symposia on Molecular and Cellular Biology, Vol. XXI), 323–340. New York: Academic Press 1982.

Webster, N.: A Brief History of Epidemic and Pestilential Diseases, Vol. 2, Section 12, 39. London: G. Woodall 1800.

Webster, R. G., Campbell, C. H., Granoff, A.: The *in vivo* production of "new" influenza viruses. I. Genetic recombination between avian and mammalian influenza viruses. Virology *44,* 317–328 (1971).

Webster, R. G., Kendal, A. P., Gerhard, W.: Analysis of antigenic drift in recently isolated influenza A (H1N1) viruses using monoclonal antibody preparations. Virology *96,* 258–264 (1979).

Webster, R. G., Laver, W. G.: Determination of the number of nonoverlapping antigenic areas on Hong Kong (H3N2) influenza virus hemagglutinin with monoclonal antibodies and the selection of variants with potential epidemiological significance. Virology *104,* 139–148 (1980).

Webster, R. G., Berton, M. T.: Analysis of antigenic drift in the hemagglutinin molecule of influenza B virus with monoclonal antibodies. J. Gen. Virol. *54,* 243–251 (1981).

Webster, R. G., Hinshaw, V. S., Bean, W. J., jr., van Wyke, K. L., Geraci, J. R., St. Aubin, D. J., Petursson, G.: Characterization of an influenza A virus from seals. Virology *113,* 712–724 (1981).

Webster, R. G., Hinshaw, V. S., Berton, M. T., Laver, W. G., Air, G.: Antigenic drift in influenza viruses and association of biological activity with the topography of the hemagglutinin molecule. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., Fox, C. F., eds.), (ICN-UCLA Symposia on Molecular and Cellular Biology, Vol. XXI), 309–322. New York: Academic Press 1981.

Webster, R. G., Hinshaw, V. S., Laver, W. G.: Selection and analysis of antigenic variants of the neuraminidase of N2 influenza viruses with monoclonal antibodies. Virology *117,* 93–104 (1982).

Webster, R. G., Laver, W. G., Air, G. M., Schild, G. C.: Molecular mechanisms of variation in influenza viruses. Nature *296,* 115–121 (1982).

Webster, R. G., Brown, L. E., Jackson, D. C.: Changes in the antigenicity of the hemagglutinin molecule of H3 influenza virus at acidic pH. Virology *126* (in press) 1983.

W. H. O. Memorandum: A revised system of nomenclature for influenza viruses. Bull. Wld. Hlth. Org. *45,* 119–124 (1971).

W. H. O. Memorandum: A revised system of nomenclature for influenza viruses. Bull. Wld. Hlth. Org. *58,* 585–591 (1980).

Wiley, D. C., Wilson, I. A., Skehel, J. J.: Structural identification of the antibody-binding sites of Hong Kong influenza hemagglutinin and their involvement in antigenic variation. Nature *289,* 373–378 (1981).

Wilson, I. A., Skehel, J. J., Wiley, D. C.: Structure of the hemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature *289,* 366–373 (1981).

Winter, G., Fields, S.: Cloning of influenza cDNA into M13: The sequence of the RNA segment encoding the A/PR/8/34 matrix protein. Nucl. Acids Res. *8,* 1965–1974 (1980).

Winter, G., Fields, S.: The structure of the gene encoding the nucleoprotein of human influenza virus A/PR/8/34. Virology *114,* 423–428 (1981).

Winter, G., Fields, S., Brownlee, G. G.: Nucleotide sequence of the hemagglutinin gene of a human influenza virus H1 subtype. Nature *292,* 72–75 (1981).

Winter, G., Fields, S., Gait, M., Brownlee, G.: The use of synthetic oligodeoxynucleotide primers in cloning and sequencing segment 8 of influenza virus (A/PR/8/34). Nucl. Acids Res. *9,* 237–245 (1981).

Wrigley, N. G., Skehel, J. J., Charlwood, P. A., Brand, C. M.: The size and shape of influenza virus neuraminidase. Virology *51,* 525–529 (1973).

Wrigley, N. G.: Electron microscopy of influenza virus. Brit. Med. Bull. *35,* 35–38 (1979).

Yewdell, J. W., Webster, R. G., Gerhard, W.: Antigenic variation in three distinct determinants of an influenza type A hemagglutinin molecule. Nature *279,* 246–248 (1979).

Young, J. F., Palese, P.: Evolution of human influenza A viruses in nature: Recombination contributes to genetic variation of H1N1 strains. Proc. Natl. Acad. Sci. U.S.A. *76,* 6547–6551 (1979).

6

Expression of Cloned Influenza Virus Genes

Mary-Jane Gething and *J. Sambrook*

I. Introduction

The cloning in bacterial plasmids of double-stranded DNA copies of the various RNA segments of the influenza virus genome has resulted in an unprecedented expansion of our knowledge about the proteins of the virus. Analysis of the DNA sequence of the cloned genes, added to the information available from earlier classical protein chemistry studies, has led to the elucidation of the amino acid sequences of all the known virus-coded proteins and to the discovery of hitherto unsuspected polypeptides; comparisons of the primary structures of the proteins coded by viruses of different subtypes has enhanced our understanding of the mechanisms of antigenic drift and shift; furthermore, knowledge of amino acid sequences has greatly simplified the interpretation of the three-dimensional structures of the external domains of the haemagglutinin (HA) and neuraminidase (NA) glycoproteins and allowed antigenic sites to be accurately mapped on the HA structure.

Recombinant DNA technology has brought another major benefit in opening the way to express cloned influenza virus genes in both prokaryotic and eukaryotic hosts. These studies have two major aims. The first is to obtain large quantities of pure HA or NA antigens efficiently and inexpensively for use as vaccines. The second is to study, in prokaryotic or eukaryotic cells, the biosynthesis, structure and function of individual wild-type or mutated influenza virus proteins. Because these proteins are naturally encoded by a negative-strand RNA genome it has not previously been feasible to manipulate their primary structures by directed alterations of the nucleotide sequences that encode them.

While the majority of expression studies in bacterial and mammalian cells have utilized cloned copies of the viral HA gene, the matrix (M) and nonstructural (NS) genes also have recently been expressed and it is certain that similar studies with the other viral genes will follow in the near future. Thus we can anticipate a rapid expansion of our knowledge of the biological properties of each protein and of its role in the life cycle of the virus.

II. Expression Vectors

Many vector systems are now available for the expression of eukaryotic genes in bacterial or mammalian cells (for reviews see Maniatis *et al.*, 1982; Gluzman, 1982). Vectors that have been used to express cloned influenza genes contain, in addition to the DNA sequences necessary for their replication in host cells, controlling elements that promote efficient transcription of the exogenous genes and translation of the resulting mRNAs.

Expression of HA and NS in *E. coli* has been achieved by inserting cDNA copies of the genes into a variety of plasmid vectors. These plasmids are double-stranded, closed circular DNA molecules (3—5 kb in length) that replicate autonomously as extrachromosomal genetic elements in *E. coli*. The plasmids used for expression experiments are designed so that the foreign genes can be inserted behind powerful prokaryotic promoters to express hybrid mRNAs that carry efficient bacterial ribosome binding sites. After transformation with these recombinant plasmids, strains of bacteria are isolated that express influenza virus sequences either constitutively or inducibly, depending on the particular promoter chosen.

The synthesis of HA and M in mammalian cells has been achieved using vectors derived from simian virus 40 (SV40). In cultured cells, SV40 undergoes a conventional replication cycle in which the expression of the viral early genes (large and small T antigens) is followed by replication of the viral DNA, synthesis of the late viral proteins and finally, assembly of progeny virus particles (Tooze, 1980). The genome of the virus is a closed, double-stranded circular molecule, some 5.2 kb in length. By replacing the coding sequences of either the early or the late SV40 genes with those of the HA or M genes, it has been possible to generate recombinants that synthesize large quantities of influenza proteins. Because these recombinants are defective, their lytic growth depends on complementation by the appropriate wild-type gene products. The system can be further adapted to provide transient expression following transfection of mammalian cells with chimeric SV40-influenza DNA molecules, or continuous, low-level expression from a chromosomally integrated gene.

The characteristics of each vector system, together with the results of the expression experiments, are described below.

III. Expression of Influenza Virus Genes in E. coli

The major requirements for efficient expression of a cloned eukaryotic gene in *E. coli* have been reviewed in Maniatis *et al.* (1982). Briefly:

(i) The coding region of the gene must not be interrupted by intervening sequences.

(ii) The gene must be placed under the control of an *E. coli* promoter that is efficiently recognized by *E. coli* RNA polymerase. However it is also desirable that a promoter be chosen whose activity can be regulated, because the constitutive synthesis of large quantities of the recombinant gene products may be toxic to *E. coli* cells (see below). Examples of promoters whose activities can be regulated are *lac*uv5, *trp*, p_L, and *omp*F.

(iii) The chimeric mRNA should be efficiently translated. Usually this is achieved by providing the recombinant mRNA with a strong bacterial ribosome binding site (which is known to include the initiation codon [AUG] and a "Shine-Dalgarno" sequence 3–9 nucleotides long located 3–11 nucleotides upstream from the initiation codon). During the construction of an expression plasmid, the foreign gene may be positioned so that [1] its own AUG is at a predetermined or experimentally established optimal distance from the bacterial Shine-Dalgarno sequence, [2] it is fused in the correct reading frame to an AUG present on the vector, or [3] it is fused in frame to the 5′ coding sequences of a bacterial gene. In the first two cases, the foreign gene will be expressed as a native, unfused protein, while in the latter case a fusion protein will be obtained. All three alternatives have been utilized for the expression of influenza virus genes (see Table 1).

(iv) The foreign protein produced in *E. coli* should be stable and not rapidly degraded by bacterial proteases. If the native protein is unstable in bacteria, expression of the gene product as part of a fusion protein may improve the yield. Alternatively, some proteins may be stabilized by synthesis in large amounts, possibly because of the formation of insoluble protein precipitates that are resistant to proteases.

Several groups have succeeded in expressing cloned HA genes of various influenza virus strains in bacteria using plasmids containing either *lac* or *trp* promoters (Emtage *et al.*, 1980; Heiland and Gething, 1981; Davis *et al.*, 1981). More recently, the NS genes also have been expressed from the *pL* promoter of bacteriophage lambda (Young *et al.*, 1983). A summary of these experiments and their results is presented in Table 1.

A. Expression of HA

Levels of expression of HA in *E. coli* vary greatly (from less than 100 to 70,000 molecules per cell). This variation is eloquent testimony to the fact that the rules governing efficient expression of eukaryotic proteins in bacteria are not yet clearly understood. The different quantities of protein synthesized by different constructs may be a reflection of the relative efficiencies of the different promoters and ribosome binding sites, or a consequence of the structure and stability of each individual hybrid mRNA or protein. Despite this variation, it is clear that under some circumstances HA can be synthesized in very large quantities in *E. coli*. 50,000 molecules per cell is equivalent before purification to a yield of approximately 6 mg from a litre culture containing 10^9 bacteria/ml. In all experiments so far reported, the HA protein has been synthesized and accumulated within the bacterial cell. In most cases, the HA has been expressed as a fusion protein lacking its own signal sequence but having additional N-terminal sequences of bacterial origin linked to the mature HA polypeptide (or to a truncated HA molecule). The additional sequences may be just a few residues (*e.g.* 7 amino acids of β-galactosidase) or a substantial portion of a bacterial protein (see Table 1.). While this may serve to ensure that the chimeric mRNA is translated efficiently and may also protect the HA from proteolysis, the additional foreign sequences may pose problems if the

Table 1. *Expression of influenza virus genes in E. coli—a summary*

Gene	Bacterial promoter	Ribosome binding site	Additional N-terminal sequence	HA sequence (amino acids)	Induction	Expression level % total protein; molecules/cell
HA (FPV)[a]	*trp*	*trpE*	7 a.a. *trpE* plus 14 a.a. from linker & non-translated sequences	–18–547	–	0.75%; ~7,500
					IAA[e]	2–3%; ~25,000
HA (A/Japan/305)[b]	*lacUV5*	*lac*	7 a.a. β-galactosidase	10–547	–	n.d.; ~3,000
HA (A/WSN/33)[c]	*trp*	*trpL*	–	1–396	–	–; <100
				1–565	–	–; <100
	trp	*trpL*	trplE protein	1–396	–	low; n.d.
	lac	*lac*	–	1–396	–	–; <100
				1–565	–	–; <100
	lac	*lac*	1006 a.a. β-galactosidase	1–396	–	1–2%; 10–20,000
					IPTG[f]	5–7%; 50–70,000
				59–211	–	1–2%; 10–20,000
					IPTG[f]	5–7%; 50–70,000
NS (A/PR/8/34)[d]	*λpL*	*cII*	–	1–230	42°	10–20%; 100–200,000

[a] Emtage *et al.* (1980)
[b] Heiland and Gething (1981)
[c] Davis *et al.* (1981)
[d] Young *et al.* (1983)
[e] 3-indolylacetic acid
[f] isopropyl-β-D-thiogalactoside

recombinant product is to be used as an immunogen. In addition, the presence of an N-terminal extension on the nascent HA molecule would perhaps preclude the activity of the HA signal sequence. This factor may be of importance, because the presence of a functional signal sequence could facilitate the efficient secretion of the recombinant HA through the bacterial membrane and thus alleviate problems of protease sensitivity and simplify protein purification. Secretion would require either that the natural HA signal which works in eukaryotic cells should also function in bacteria, or that a bacterial signal be fused to the N-terminus of mature HA. Preliminary results indicate that the expression of an HA protein containing a functional hydrophobic signal sequence and a C-terminal hydrophobic anchor peptide may be toxic for bacterial growth (Heiland and Gething, 1981). However, it seems likely that this toxicity is a consequence of the accumulation in the bacterial membrane of a large amount of foreign protein. This problem can be surmounted either by deletion of the DNA sequence coding for the anchor, or by propagation and amplification of the recombinant plasmids in bacteria under conditions in which the promoter is repressed (Gething, unpublished results). Expression of HA can then be obtained after addition of the appropriate inducer.

While some interesting scientific problems such as the efficacy of eukaryotic signal sequences in bacteria can be addressed by prokaryotic expression experiments, the major thrust of this work is towards production in bacteria of antigens for use as vaccines. Although none of the HAs that have been expressed in bacteria are from strains of virus that are currently causing disease, the results obtained so far are encouraging. However, several problems remain to be solved before the utility of HA produced in bacteria can be proven. The protein is usually obtained in an insoluble and intractable form which makes its purification in a native configuration difficult. This may provide an explanation for the poor immunogenicity of the recombinant HA in comparison to the viral protein (D. P. Nayak, personal communication). Lack of glycosylation by the bacterial host may also affect the stability and immunogenicity of the protein. Although the recombinant HA can be recognized and precipitated by polyspecific anti-HA antisera, it remains to be shown whether all of the antigenic sites of the HA molecule are displayed on this protein. Unless the majority of the antigenic epitopes can be expressed in bacteria, the prospects for useful vaccine production are diminished, since all of these sites undergo change during antigenic drift or shift. Despite these obstacles, the prospect of using recombinant DNA technology to keep apace of the rapid evolution of the HA sequence and to produce inexpensive vaccines remains extremely attractive.

B. Expression of NS1 Protein

The NS gene from the A/PR/8/34 strain of influenza virus has been expressed at very high levels in *E. coli* using the p_L promoter of bacteriophage lambda (Young *et al.*, 1983). The resulting NS1 protein has been purified and used to raise an antiserum that recognises the protein synthesized in eukaryotic cells. The satisfactory

immunogenicity of this protein may reflect the fact that, unlike HA, NS1 is not a glycoprotein and the protein produced in bacteria may therefore more closely resemble the NS molecule synthesized in eukaryotic cells.

IV. Expression of Influenza Virus Genes in Eukaryotic Cells

A. Expression of HA in Simian Cells Using Recombinant SV40 Viral Vectors

The HA genes from three different subtypes of influenza virus (A/WSN/33 [H0N1], Hartman *et al.*, 1982; A/Japan/305/57 [H2N2], Gething and Sambrook, 1981; A/Udorn/72 [H3N2], Sveda and Lai, 1981) have been inserted into viral vectors derived from SV40 and expressed in simian cells. The construction of the recombinant genomes, their introduction into eukaryotic cells and the characteristics of the HA proteins expressed are described below.

Because the HA gene occurs naturally in the form of a negative-strand RNA, the cloned DNA copy contains none of the controlling elements, such as promoters and enhancers, that are required for efficient transcription of conventional DNA genes. Thus, the SV40 vectors have been designed so that the HA coding region is placed under the control of either the early SV40 promoter/enhancer or the SV40 late promoter(s). The untranslated 3′ region of the HA gene contains a sequence that is closely related to the consensus polyadenylation signal for eukaryotic genes and that is believed to function as such a signal during natural infections with influenza virus. Initially, however, it was not clear that this signal would work properly in the form of double-stranded DNA. Therefore, to ensure that transcripts of the HA gene would be terminated efficiently, the HA sequences were always inserted upstream of a poly(A)-addition signal in the SV40 sequence. Furthermore, at the time that these experiments were undertaken, there were indications from other workers that foreign genes inserted into SV40 recombinant vectors were expressed more efficiently when splice donors/acceptors were present in the primary transcript (Mulligan *et al.*, 1979; Hamer and Leder, 1979). Consequently, some of the vectors were constructed so that either the natural SV40 splicing signals were retained in 5′ non-coding sequences of the hybrid transcript, or an exogenous intron sequence was incorporated into the 3′ non-coding sequence of the recombinant mRNA. The efficacy of these signals is discussed below.

A recombinant SV40 virus will replicate efficiently in eukaryotic cells only if (i) the host cells are permissive for viral DNA synthesis, (ii) the vector contains a functional origin of replication, (iii) there is a supply of large T antigen sufficient to initiate successive rounds of SV40 DNA synthesis, and (iv) a supply of the SV40 capsid proteins VP1, VP2 and VP3 is available for production of infectious virions containing the recombinant genome (Gluzman, 1981; Tooze, 1980). Because the late replacement recombinants contain a functional origin of DNA replication and an intact set of SV40 early genes, their genomes will replicate efficiently in permissive simian cells. However the late genes of SV40 have been deleted and production of

infectious virions containing the recombinant genome therefore requires that SV40 capsid proteins be supplied by a complementing helper virus—for example an SV40 mutant carrying a defect in the early region of the viral genome. On the other hand, the early replacement recombinants contain an intact set of late genes, but lack the SV40 early genes, so that their genomes cannot replicate in simian cells unless functional T antigen is supplied. This can be done by using as a permissive host the COS-1 line of SV40-transformed simian cells that carry an endogenously expressed copy of the SV40 T antigen gene (Gluzman, 1981). Finally, efficient packaging of recombinant genomes into SV40 capsids requires that the genome size remains within 70—105% of that of the wild-type virus (Tooze, 1980). This restraint can be accomodated by choice of the sites between which the foreign gene is to be inserted, and by use of viable deletion mutants of SV40 virus which provide additional packaging capacity.

B. Preparation of the HA Genes

When the various HA genes were originally cloned in plasmids, the coding sequences were preceded by approximately 40 nucleotides of untranslated sequence as well as the homopolymeric dG : dC and/or dA : dT tails that had been employed during the cloning procedures (Lai *et al.*, 1980; Gething *et al.*, 1980; Davis *et al.*, 1980). These 5′ sequences were removed from the Japan and WSN genes by treatment with the double-stranded exonuclease, Bal31. In the case of the Japan gene, the DNA sequences of the resected clones were analysed and one clone was chosen in which the first nucleotide of HA sequence was the A of the initiating AUG (Gething and Sambrook, 1981). Although the Bal31-treated WSN HA clones were not sequenced, restriction endonuclease analysis was used to identify clones that appeared to have lost homopolymeric tails and untranslated sequences (Hartman *et al.*, 1982). By contrast, the Udorn HA gene that was inserted into the SV40 late region still contained dG : dC tails and untranslated nucleotides (Sveda and Lai, 1981).

C. Construction of Recombinant Genomes

(i) *Vectors in which the HA coding sequence replaces the SV40 early genes.* As mentioned above, the cloned Japan HA had been reconstructed so that the A residue immediately following a HindIII linker began the initiating AUG of the protein. This allowed the gene to be inserted into the early region of the SV40 genome (between the HindIII site at nucleotide 5171 and the BclI site at nucleotide 2770), so that this initiation codon was almost exactly in the position normally occupied by the initiation codon of SV40 large and small T antigens (Gething and Sambrook, 1981). The coding sequences of the HA gene lie upstream of the SV40 early transcript polyadenylation site at nucleotide 2587. The HA gene should therefore be transcribed under the control of the upstream SV40 promoter/enhancer to yield an mRNA that (i) initiates at the usual site, (ii) contains a 5′ untranslated region that apart from the last 5 nucleotides is identical to that of the conventional SV40 early mRNAs, (iii)

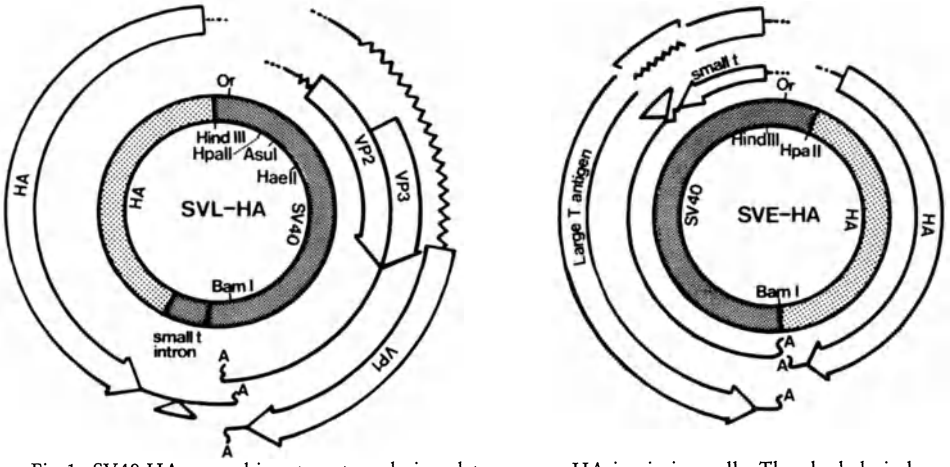


Fig. 1. SV40-HA recombinant vectors designed to express HA in simian cells. The shaded circles represent the double-stranded circular DNA genomes of the recombinant viruses. The protein-coding sequences are indicated by blocked arrows. Untranslated parts of the mRNA are shown as solid lines; dots indicate uncertainty about the exact position of the 5′ end of the mRNA; zigzag lines show segments spliced out; a wavy line with an A illustrates the 3′-terminal poly(A) tract

carries a coding region specifying an HA molecule that is identical to that of the natural form of the molecule, and (iv) terminates at either the poly(A)-addition site within the HA gene or at the SV40 site further downstream. However, the recombinant HA mRNA does not contain any splicing signals. To determine if the provision of splice donor and acceptor sites would increase the efficiency of expression of the HA gene, a second recombinant genome was constructed which contained, downstream from the HA sequences, a segment of DNA containing the intron of the SV40 gene coding for small t antigen. This recombinant is shown in Fig. 1.

(ii) *Vectors in which the HA coding sequences replace the SV40 late genes.* The promoter(s) of transcription of the SV40 late region have not been defined. The various late replacement recombinants have been constructed such that the HA genes are inserted either into the untranslated 5′ region of the SV40 late mRNAs or into the coding sequence of VP2. In all cases, the HA coding sequences are inserted upstream of the polyadenylation signal for late transcripts (nucleotide 2674). The various constructs differ in the provision of splice signals in the hybrid transcript.

To place the Japan HA coding sequences under the control of the SV40 late promoter(s), the gene was inserted into the late region of the SV40 genome between the HpaII site at nucleotide 346 and the Bam HI site at nucleotide 2533 (Fig. 1). The HA coding sequences are joined to the noncoding sequences that would normally be transcribed into the untranslated 5′ region of SV40 mRNAs (Gething and Sambrook, 1981). At late times during lytic infection of permissive simian cells, this recombinant genome should express an unspliced hybrid mRNA consisting of untranslated SV40 sequences at its 5′ end and an intact set of HA coding sequences. Polyadenylation could occur either at the site within the 3′ untranslated sequences of the HA gene or at the normal site for SV40 late mRNAs at nucleotide 2674.

Translation of this RNA should begin at the first AUG, which is located at the beginning of the HA coding sequences.

A very similar recombinant, differing only in the addition of a Bam HI linker at the HpaII site of SV40, has been constructed to express the WSN HA gene (Hartman *et al.*, 1982). This gene has also been expressed from another recombinant in which the HA coding sequences have been inserted following the AsuI site which lies at nucleotide 557 in the SV40 sequence and immediately precedes the initiation codon for VP2. The hybrid late transcript expressed from this recombinant genome has the same splice donor and acceptor sites in its 5′ untranslated sequences as are used in the maturation of the natural SV40 19S late mRNA. Once again, the HA initiation codon is the first AUG that occurs in the mRNA.

All of the previously described recombinants have the HA genes inserted into non-coding regions of SV40. Another late replacement vector contains the Udorn HA coding sequences inserted into the SV40 genome between the HaeII site at nucleotide 832 and the Bam HI site at nucleotide 2533 (Sveda and Lai, 1981). This places the HA AUG within the DNA sequence coding for VP2, so that it is not the first AUG in the hybrid transcript. A similar situation is found in wild-type SV40, where the AUG of VP3 lies within the coding sequences of VP2. The SV40-HA hybrid transcript has the same 5′ untranslated sequences and splice donor/acceptor sites as the WSN construct described above.

D. Introduction of Recombinant Genomes into Simian Cells and Production of Virus Stocks

Constructions of the HA-SV40 recombinants were carried out in bacterial plasmids of various sorts. The Japan HA-SV40 recombinant genomes were also propagated in bacterial plasmids until their structures had been verified by restriction endonuclease digestion and analysis of the DNA sequence across the junctions between the SV40 and HA genes (Gething and Sambrook, 1981). The hybrid genomes were then excised from the plasmids (by cleavage with Bam HI), purified by gel electrophoresis and ligated at low concentrations of DNA to generate closed circular molecules that were then introduced into the appropriate host cells using DEAE-dextran as facilitator (McCutchan and Pagano, 1968). The early-replacement recombinants were introduced into COS-1 cells in the absence of helper DNA; the late-replacement recombinant was introduced into CV-1 cells together with an equal amount of the DNA of the SV40 early deletion mutant dl1055 (Pipas *et al.*, 1980).

By contrast, the DNA fragments used in the construction of the late-replacement recombinants containing the WSN or Udorn HA genes were purified separately and then ligated together *in vitro* (Sveda and Lai, 1981; Hartman *et al.*, 1982). The ligation mixtures were transfected directly into AGMK cells without prior analysis or cloning of the recombinant genomes. The cultures were subsequently superinfected with the complementing helper viruses tsA28 or pSVL5 and recombinant viruses were recovered either as mixed stocks or from individual plaques.

Only a small proportion of cultured mammalian cells are competent at any one time to take up DNA, so that very few cells of the population are infected initially. However, during the next several days, each of these cells undergoes a lytic infection and produces over a million virus particles that are able to spread into neighbouring cells and infect them with high efficiency. Usually, the lysate obtained from the first set of infected cells needs to be passed serially once or twice to obtain a high titre virus stock. In the case of the early-replacement recombinants, this virus stock consists exclusively of recombinant genomes encapsidated in SV40 coat proteins; in the case of the late-replacement recombinants, it consists of approximately equal numbers of helper virus particles and recombinants. Both types of virus stocks can be used to infect permissive cells with high efficiency and to induce lytic cycles of virus growth. During such infections, the viral genomes are transported to the nucleus, where they are liberated from their capsids. The early promoter soon becomes active and genes under its control are expressed. By 12 hours after infection, viral DNA replication is under way and genes begin to be expressed from the late promoter with high efficiency. By 36–48 hours infection, the newly-synthesized viral genomes begin to be assembled into progeny virus particles—a process that continues for another 24 hours, when the cells detach from their substrate and die.

E. Analysis of the HA Expressed from the SV40-HA Recombinant Vectors

All of the recombinant vectors described above express an HA glycoprotein that is indistinguishable in all respects tested from HA synthesized during influenza virus infections. To establish that the protein was authentic in its structure, antigenicity and biological activities, a series of assays were performed on cells infected with the various recombinants. These included immunoprecipitation, immunofluorescence, haemagglutination and cell fusion assays (Fig. 2). These experiments have been described in detail previously (Gething and Sambrook, 1981; Sveda *et al.*, 1981; Hartman *et al.*, 1982; White *et al.*, 1982) and will be discussed only briefly here.

Immunoprecipitation: extracts of cells infected with the recombinant viruses contained a protein that could be labelled either with ^{35}S-methionine or with radioactive sugar precursors and that was specifically precipitated by anti-HA sera. The protein was indistinguishable in size from authentic, glycosylated HA precipitated from extracts of cells infected with influenza virus. Cells infected with the SVEHA3 recombinant virus stock contained so much of this protein that it was not necessary to use immunoprecipitation to detect it; HA could be seen either as a band stained with Coomassie blue after the extracts of infected cells had been analysed by SDS-PAGE, or as a prominent radioactive species when extracts were prepared from infected cells labelled with ^{35}S-methionine. Treatment of the infected cells with tunicamycin resulted in the production of a smaller non-glycosylated protein that was identical in size to non-glycosylated authentic HA.

Immunofluorescence: cells infected with the SV40-HA recombinants displayed bright, cytoplasmic fluorescence, with especially intense staining of the Golgi

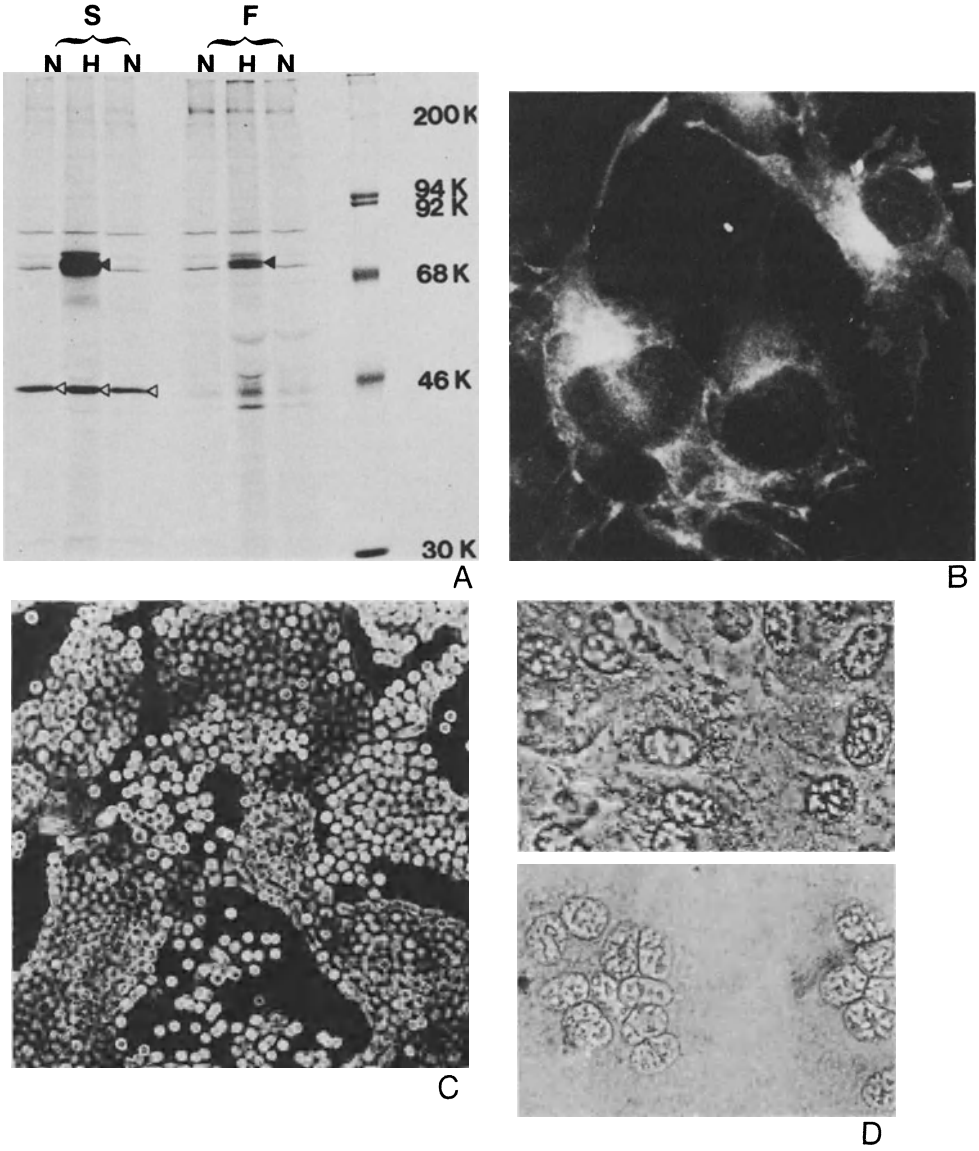


Fig. 2. Analysis of the HA expressed from SV40-HA recombinant vectors. The details of the experimental protocols have been described previously (Gething and Sambrook, 1981; White *et al.*, 1982). Panel *A:* Immunoprecipitation, using rabbit anti-HA (HA) *(H)* or non-immune *(N)* sera, of extracts of CV-1 cells infected with SVEHA3 *(S)* or influenza virus *(F)*. Panel *B:* immunofluorescent antibody staining of HA in CV-1 cells infected with SVLHA8. Panel *C:* adsorbtion of guinea pig erythrocytes to a monolayer of CV-1 cells infected with SVEHA3. Panel *D:* Low-pH induced fusion of a monolayer infected with SVEHA3. Fusion occurs after transient incubation at pH5 (both panels) when the infected cells have previously been treated with trypsin (lower panel only) to proteolytically activate the cell surface HA

Table 2. *Quantitation of the expression of HA in extracts of simian cells infected with SV40-HA recombinant viruses. Monolayers of CV-1 or COS-1 cells were infected as described previously. Sixty hours after infection, cell extracts were prepared and the amount of HA was estimated by solid phase radioimmune assay (Gething and Sambrook, 1982)*

Recombinant virus	μg HA per 106 cells	Av. no. molecules per cell	% cells infected	Molecules HA per infected cell
SVEHA3	57	5.7×10^8	95	6.0×10^8
SVLHA8 (intron$^+$)	0.08	8.0×10^5	5	1.6×10^6
SVLHA10 (intron$^-$)	14	1.4×10^8	95	1.5×10^8

apparatus. The surface of the cells also stained specifically, with a uniform, dimmer fluorescence. The distribution of fluorescence in cells infected with influenza virus was similar to that displayed by recombinant infected cells, but of lower intensity.

Haemagglutination and erythrocyte binding to infected cells: guinea pig red blood cells could be agglutinated by extracts of cells infected with the recombinant viruses. Intact monolayers of infected cells adsorbed a dense carpet of erythrocytes onto their surfaces.

Cell-cell fusion: CV-1 cells infected with the SVEHA3 recombinant could be fused to form giant polykaryons if the monolayers were first treated with low levels of trypsin to cleave HA0 to HA1 and HA2 and then exposed to transient low pH (White *et al.*, 1982). The pH threshold at which the fusion occurs is identical to that observed for the low pH-induced fusion of cells infected with the Japan strain of influenza virus.

F. Quantitation of the Amounts of HA Expressed from the Recombinant Genomes

The amounts of Japan HA expressed from the early- and late-replacement recombinants has been quantitated using a solid phase radioimmune assay (Gething and Sambrook, 1981). It was found that the amount of HA detected increased as the course of the infection proceeded. By 62 hours, when the lytic infection was in its terminal phase, cells infected with the late-replacement recombinant, SVEHA3, contained approximately 6×10^8 molecules of HA per cell, which corresponds to the production of about 200 μg of HA per 10 cm dish of cells. For comparison, a simian cell at a late stage during infection with influenza virus will contain about 5×10^7 molecules of HA per cell. Cells infected with the intron-minus early-replacement recombinant, SVLHA10, also produced HA with high efficiency (see Table 2). These results show that the cloned HA gene can be expressed at very high levels when it is coupled to either the early or the late SV40 promoters. However, SVLHA8(intron$^+$)-infected cells produced much lower amounts of HA antigen (Table 2). The reason for this decreased production of HA will be discussed in the next section.

G. Effect of Intervening Sequences in the Recombinant HA Transcript on the Level of Expression of HA Protein

The effect of the presence of splicing signals in the recombinant transcript on the expression of HA and the stability of the genome is shown in Table 3. It was clear from these studies that the splicing of mRNA is not required for expression of the HA gene from either the early or late replacement recombinants. A possible correlation between the presence of an intron and instability is best illustrated by comparison of the two early replacement recombinants which are identical in every respect except for the addition of a DNA segment containing the SV40 small t antigen intron between the HA coding sequences and the poly(A)-addition site (Gething and Sambrook, 1982 a). When the abilities of the two recombinants to express HA were compared, it was found that while 95% of the cells infected with the intron$^-$ virus expressed HA, only 5% of those infected with the intron$^+$ virus did so. Analysis of the recombinant DNA molecules isolated from cells infected with sequential passages of the intron$^+$ virus showed a rapid increase in the proportion of rearranged genomes. By the fourth and fifth passages, the stock was essentially non-infectious and contained a preponderance of molecules shorter than that of the original construct. Cloning and sequencing of a number of the altered molecules revealed that many forms of rearrangements had occurred, including simple deletions within the HA gene and across both of the SV40-HA borders, and deletions of HA sequences together with insertions or inversions of SV40 DNA. Although the genome of the intron$^+$ recombinant appears to be particularly vulnerable to such rearrangements, the intron$^-$ genome is completely stable, as is the late-replacement recombinant SVEHA3, which also lacks any splicing signals in the HA transcript. It might be tempting to argue that the presence of the intron is actually detrimental to the expression of the HA mRNA. However, the rapid accumulation of rearrangements in the recombinant genome suggests that the deleterious effects of the intron are mediated more at the level of DNA than of RNA. Why genomic instability should be correlated with the presence of the intron remains to be elucidated.

Table 3. *Correlation between the presence of splicing signals, the expression of HA and the stability of SV-HA recombinant genomes*

Recombinant	HA gene	Intron	Genome rearrangement	Expression level
(Late)				
SVEHA3	Japan	−	−	+++
SVHA1	WSN	−	?	+
SVHA2	WSN	+	?	+
SVHA3	WSN	+	?	+
HA-SV40-8	Udorn	+	+	+
(Early)				
SVEHA10	Japan	−	−	+++
SVEHA8	Japan	+	++	+

H. Expression of the Matrix Gene in Simian Cells Using an SV40-M Recombinant Virus

A cloned copy of influenza virion RNA segment 7, which codes for the viral M proteins, has been inserted into the late region of SV40 between the HpaII site at nucleotide 346 (which had previously been converted to a BamHI site) and the BamHI site at nucleotide 2533 (Lamb and Lai, 1982). The recombinant SV40-M genome was introduced into AGMK cells using DEAE-dextran together with SV40 tsA28 helper virus and a cloned recombinant virus stock was prepared in CV-1 cells. Analysis of M-specific mRNAs produced in cells infected with this stock showed that both spliced and unspliced hybrid mRNAs having 5′ and 3′ non-coding sequences derived from SV40 were synthesized. The splice sites utilized to form the mRNA were identical to those used to form $mRNA_3$ during *in vivo* influenza virus transcription (Lamb *et al.*, 1981). However, a spliced product equivalent to the authentic $mRNA_2$ was not observed and the relative abundance of the interrupted and colinear RNA species varied in cells infected with influenza virus or the SV40-M recombinant. These differences may be due to effects of the SV40 sequences at the 5′ and 3′ ends of the transcript. Immunoprecipitation of the polypeptides synthesized in SV40-M-infected cells with influenza virus-specific antisera and tryptic peptide mapping demonstrated a polypeptide with exactly the same mobility and peptide composition as influenza virus M_1 protein. Such a protein could only be translated from the unspliced hybrid mRNA containing the complete coding sequence of the M_1 protein.

I. Transient Expression of Cloned Genes in COS-1 Cells

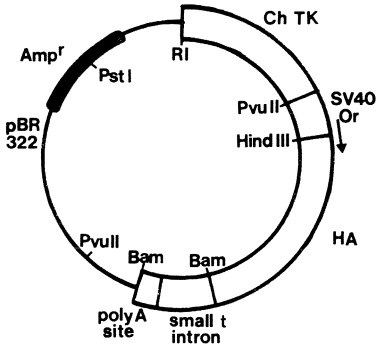
The experiments described above involve the production and use of virus stocks containing recombinant genomes. Such stocks are of great value when efficient infection of cells is required for high level expression of the protein products. However, several passages through cells are required before high titre stocks are developed and a more rapid screening procedure is necessary when a large number of recombinants must be checked for expression—for example, during the analysis of the phenotypes of mutated genes. Fortunately, it is possible to study, using COS-1 cells, the transient production of RNA and protein from recombinant genomes which contain an SV40 origin of replication (Gluzman, 1980; Mellon *et al.*, 1981). When such DNA molecules are transfected into COS-1 cells (which contain an endogenous copy of the SV40 large T antigen gene) they are amplified to several thousand copies per cell, enabling the expression of the genome to be analysed.

The Japan HA gene carried in SV40 recombinant genomes can be expressed transiently when it is transfected into COS-1 cells using DEAE-dextran. Fourty-eight hours after transfection, 5% to 20% of the cells express HA, as judged by immunofluorescence or erythrocyte binding, and the yield of HA as measured by radioimmune assay is 1% to 5% of the levels obtained when all the cells in the culture are infected with the recombinant virus stock in the conventional manner (Gething, unpublished results).

J. Continuous Expression of HA from Genes Integrated into the Chromosomes of Eukaryotic Cells

The eukaryotic vector systems described above provide only short term expression of passenger genes either because the host cells die as a consequence of the lytic infection or because the chimeric genomes are only transiently present in the cells. Cell lines that continuously express individual cloned influenza virus genes would provide a more convenient source of protein. Furthermore, by exploiting the fusion activity of HA, they may lead to the development of systems for delivering exogenous genes and proteins into cells. Finally, they open the study of the regulation of cytotoxic T cell responses to virus-specific surface antigens.

When monolayers of eukaryotic cells are treated with DNA co-precipitated with calcium phosphate (Wigler *et al.*, 1979), a small proportion of the cells take up the DNA and integrate it into their chromosomes. Because this event is so rare, the DNA of interest is usually introduced into the cells together with a gene which will confer a selective advantage upon those cells which receive it. For example, the thymidine kinase (tk) gene will confer resistance to HAT medium upon tk^- recipient cells. Initially, the Japan HA gene was introduced into murine $LMtk^-$ cells in a recombinant plasmid containing (i) the β-lactamase gene, (ii) the chicken thymidine kinase gene (Perucho *et al.*, 1980) and (iii) the HA gene inserted between the SV40 virus early promoter and SV40 early sequences containing RNA processing signals (Fig. 3). Subsequent experiments involved co-transformation of cells with two separate plasmids; one (either pSVEHA8 or pSVLHA10 [Fig. 1]) contained the HA gene under the control of the SV40 early promoter while the other contained the chicken tk gene (p0D3, D. Hanahan, unpublished). Either murine $LMtk^-$ or human $143tk^-$ cells were used as recipients in these studies (Gething and Sambrook, unpublished results). In all cases, tk^+ clones were selected in HAT



pTKSVHA-1

Fig. 3. A recombinant plasmid designed for the constitutive expression of HA in continuous lines of murine or human cells

medium, grown into mass culture and analysed (i) for the presence of integrated HA DNA by Southern blotting, (ii) for HA mRNA by Northern transfer and (iii) for HA protein by radioimmune assay. Whereas all the tk^+ clones tested had intact, integrated HA genes, and some contained mRNAs corresponding in length to authentic HA mRNA from influenza virus infected cells, the levels of mRNA expression varied widely. The production of HA protein correlated with levels of mRNA and ranged from about 5000 molecules/cell to as much as 100,000 molecules/cell. These high-producer lines continued to express HA through several subclonings and continuous culture for a period of 8 months. Although agglutination of erythrocytes to these lines was not observed, evidence that the HA is expressed at the surface of the murine cells was provided by H2-restricted lysis of the cells by Japan HA-specific cytotoxic T cells (T. Braciale, personal communication). These results show that it is possible to obtain permanent cell lines that synthetize HA constitutively. However, the amounts of HA are relatively small, presumably because the number of integrated copies of the HA gene is low (1 to 2 copies per cell). In addition, the possibility exists that accumulation of large amounts of HA on its surface is deleterious to a cell, so that there is a constant selection pressure in favour of "underproducers". Both to distinguish these possibilities and to isolate permanent cell lines that produce greater amounts of HA it will be necessary (i) to amplify the number of HA genes in the cell and (ii) to reduce the potential deleterious effects of cell-associated HA by introducing the anchor-minus mutant HA gene (see below) that is secreted from the cell.

V. Analysis of the Expression of Mutant HA Proteins

It is clear from these studies that recombinant viruses can be constructed which express from the cloned HA genes large quantities of HA proteins that appear normal in all respects; their molecular weights are indistinguishable from that of the authentic protein and they are displayed on the infected cell's surface in a glycosylated form that is both antigenically and biologically active. Thus, a protein that is normally encoded by a negative-strand RNA genome can be expressed in copious amounts when double-stranded DNA copies of its coding sequences are harnessed to strong SV40 promoters. It is therefore possible to begin to think of using site-directed mutagenesis to dissect those parts of the HA molecule important to its structure, function and biosynthesis which have until now been inaccessible to genetic analysis.

The first set of HA mutants that have been analysed by expression in SV40-HA recombinants have deletions in the sequences coding for the N- and C-terminal hydrophobic regions of the protein. These regions play a crucial role in the transport of HA within the infected cell. The N-terminal signal sequence is required for the vectorial transfer of the nascent polypeptide across the membrane of the endoplasmic reticulum: the C-terminal hydrophobic sequence anchors the completed protein in the lipid bilayer of the cell membrane or the virus envelope.

A. The Signal-Minus HA is a Non-Glycosylated, Intracellular Protein

Preparation of the signal-deleted mutant of the Japan HA gene utilized a plasmid (pJHB29) from which all of the 5′ nucleotides coding for the signal had been removed using Bal31 exonuclease. This truncated HA gene was used to replace the wild-type Japan HA gene in the late expression vector, so that the HA initiation codon was fused in phase to the codon for the first amino acid of the mature HA polypeptide (Gething and Sambrook, 1982 b). The remainder of the HA coding sequences were unaltered.

When simian cells were infected with a virus stock containing the mutant recombinant, a variant, unglycosylated HA protein was detected by immunoprecipitation. The properties of this mutant HA protein are summarized in Table 4. Its molecular weight (M_r=61,000) was identical to that of the non-glycosylated form of HA synthesized either *in vivo* in the presence of tunicamycin or *in vitro* by translation in a reticulocyte lysate. This result is consistent with the synthesis of the signal-minus HA as a cytoplasmic protein on free polyribosomes. Lacking a signal sequence, the nascent polypeptide cannot be translocated through the membrane of the endoplasmic reticulum and therefore is never exposed to the glycosylation enzymes that reside on the luminal side of the membrane. Unlike the wild-type HA, which is expressed very efficiently from the SVEHA3 recombinant (6×10^8 molecules/cell), the signal-minus mutant is expressed at a relatively low level (10^6 molecules/cell). Control experiments showed this poor yield to be due neither to a failure of the mutant virus to infect CV-1 cells, nor to a deficiency in replication of the recombinant viral genome. Rather, it seems to stem from a combination of factors, including instability of both the mutant mRNA and the variant protein. These results clearly confirm that the signal sequence is necessary for correct transport of the nascent polypeptide. Further experiments utilizing mutants containing single nucleotide changes in the DNA sequence encoding the hydrophobic

Table 4. *Comparison of the properties of wild-type and mutant HAs produced from SV40-HA vectors*

		SVEHA3 Wild-type	SVEHA20-A$^-$ Anchor-minus	SVEHA7-S$^-$ Signal-minus
Location		Cell surface	Secreted to medium	Intracellular
M_r (subunit)		75K	72K	63K
Structure		Trimer	Trimer	?
Glycosylation		+	+	−
Red cell binding to infected cells		+	−	−
Molecules/cell (60 hours p.i.)		6×10^8	1×10^9	1×10^6
Amount produced per 10 cm dish (60 hours p.i.)	Extract	200 μg	28 μg	0.5 μg
	Medium	0	300 μg	0

peptide should elucidate the role of particular amino acids in signal function and determine the effect of the nucleotide sequence in this region on the stability and efficiency of translation of the HA mRNA.

B. Removal of the C-Terminal Hydrophobic Sequence Converts HA into a Secreted Protein

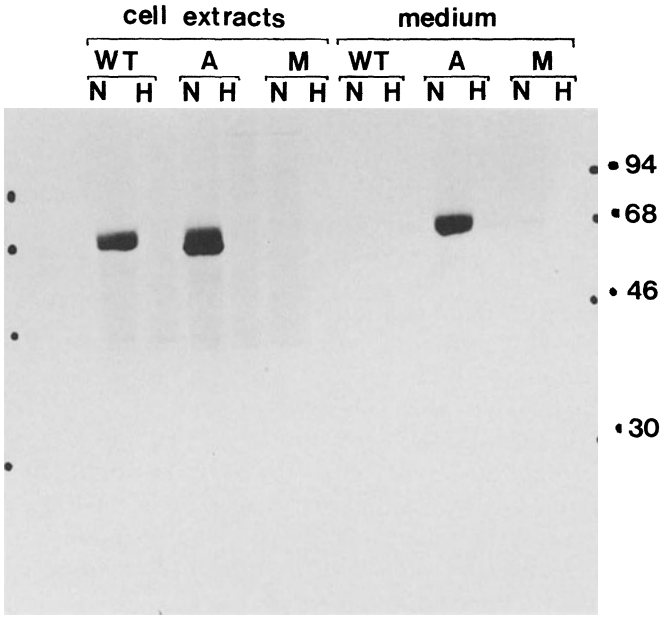
HA is normally anchored in the plasma membrane of the infected cell or in the envelope of the influenza virion by a tract of approximately 24—27 amino acids that are coded by the distal sequences of the HA gene (Gething *et al.*, 1980). These sequences have been removed from the Japan (Gething and Sambrook, 1982 b) and Udorn (Sveda *et al.*, 1982) HA genes and the expression of these deletion mutants from SV40-HA vectors has been analysed.

The construction of the Japan anchor-minus mutant utilized a plasmid (p2G10) which was originally cloned lacking sequences corresponding to the C-terminal anchor (Gething *et al.*, 1980). This truncated gene replaced the full length HA gene in the late replacement recombinant to create a new recombinant that differs from its parent in only one significant respect; it codes for a protein which lacks the 38 amino acids that are normally found at the C-terminus of HA. Substituting for them is a stretch of 11 amino acids, largely polar in nature, that are encoded by the dG : dC homopolymer tail, a synthetic Bam HI linker and a short sequence of SV40 DNA (Gething and Sambrook, 1982 b).

The Udorn deletion mutants were constructed by linearizing the genome of the late replacement recombinant with Bam HI to cut the Udorn HA sequence just before the nucleotides coding for the hydrophobic C-terminus. After S1 nuclease treatment to remove the cohesive ends, the linear molecules were transfected into simian cells and virus stocks containing deleted genomes were generated. The sizes of the deletions ranged from 40 to 174 amino acids. Substituting for these residues were short stretches of 6 to 25 amino acids derived from previously non-coding HA or SV40 sequences (Sveda *et al.*, 1982).

These anchor-minus mutant HAs are synthesized as efficiently from the recombinant genomes as the wild-type proteins. However, instead of being entirely cell-associated (in the Golgi or on the cell surface), the mutant HAs are secreted efficiently into the medium (Fig. 4). The most straightforward explanation of the mutants' properties is that removal of the C-terminal hydrophobic sequences results in loss of the anchoring function, so that the nascent polypeptide, instead of remaining attached to the luminal face of the rough endoplasmic reticulum, passes completely through the membrane. Once free in the lumen, the mutant HA is treated by the cell as if it were an authentic secretory protein, and is discharged into the medium.

The expression of the Japan mutant has been quantitated by radioimmune assay and the results are shown in Table 4, together with a summary of its properties (Gething and Sambrook, 1982 b). The mutant HA is fully glycosylated and, like the wild-type protein (Wiley *et al.*, 1977), it is assembled as a trimer of HA subunits. The



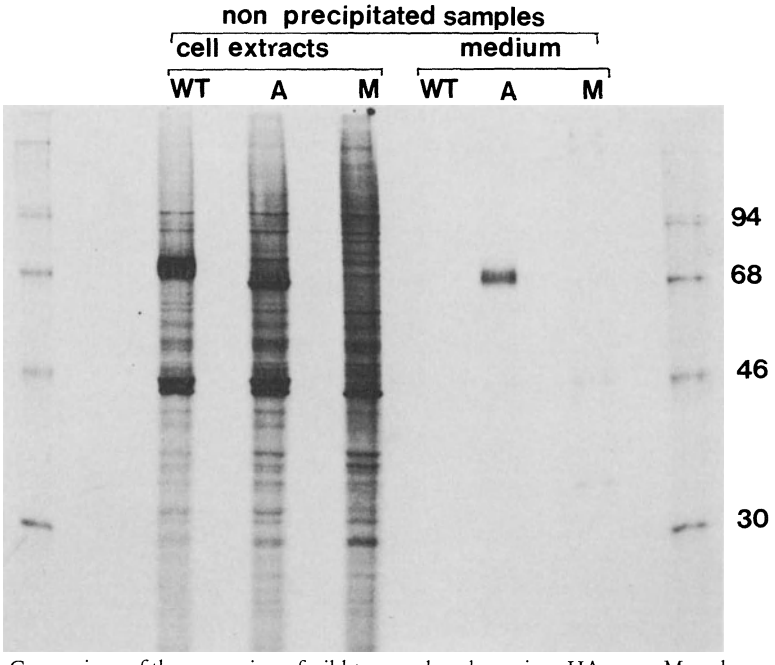


Fig. 4. Comparison of the expression of wild-type and anchor-minus HA genes. Monolayers of CV-1 cells were infected with SVEHA3 (WT), or SVEHA20-A$^-$ (A$^-$) viruses as described previously; 48 hours later the cells were labeled for 1 hour with ^{35}S-methionine and then the media above the cells were collected and cell extracts were prepared (Gething and Sambrook, 1981). Aliquots of these samples were immunoprecipitated using rabbit anti-HA serum *(H)* or normal rabbit serum *(N)*. The precipitated proteins (upper panel) together with non-precipitated samples (lower panel) were separated by SDS-PAGE and autoradiographed

reduction of approximately 3000 in its apparent molecular weight is consistent with the extent of deletion in its gene. Cells infected with the mutant do not display erythrocyte binding or pH-dependent cell fusion because the anchor-minus HA cannot form the necessary bridge between the lipid bilayer of the infected cell and that of the erythrocyte or neighbouring cell. Finally, the secreted HA has no haemagglutinating activity since, lacking the C-terminal hydrophobic region, the protein cannot form multivalent complexes.

The accumulation of large amounts of secreted HA in the medium above infected cells greatly facilitates the purification of the protein. Once the infection is well established (about 20 hours post-infection), serum-containing medium is no longer required for the continued production and secretion of HA. Thus, at 60 hours after infection, 40 hours after changing to serum-free medium, some 300 μg of essentially pure Japan HA can be harvested from a single 10 cm plate of cells infected with the anchor-minus mutant. For the first time we can purify with ease large quantities of the uncleaved form of the HA spike. Studies are underway to crystallize this material as a first step in the elucidation of the three-dimensional structure of the precursor HA molecule (Wiley, personal communication). In addition, the possibility arises of using the HA produced in tissue culture as an immunogen for vaccine production.

The secreted form of HA appears to differ from the membrane-bound wild-type protein in only one major respect—its rate of glycosylation. The addition of carbohydrate to proteins takes place in two stages. As the nascent protein appears on the luminal side of the rough ER, preformed mannose-rich oligosaccharides are transferred from a lipid carrier to certain asparagine residues (Hubbard and Robbins, 1979; Neuberger *et al.*, 1972). This initial co-translational "core" glycosylation is but the first step in an elaborate program of reactions which take place both in the rough ER and later in the Golgi apparatus, in which sugars are trimmed from and added to the nascent protein (Hubbard and Robbins, 1979; Tabas and Kornfeld, 1979). During the biosynthesis of HA, the transition from the core-glycosylated molecule to the completed molecule can be visualized by SDS-PAGE analysis of protein labelled with ^{35}S-methionine or tritiated sugar precursors in pulse-chase experiments. The final composition of the oligosaccharide side chains on the completed wild-type and anchor-minus proteins appears to be very similar. However, by contrast to the membrane-bound wild-type protein, which is glycosylated rapidly and relatively synchronously, the population of secretory molecules becomes terminally glycosylated over a very protracted period. It is possible that this difference in some measure reflects the relative efficiency with which membrane bound proteins and luminal proteins are sequestered into the transport vesicles that travel from the rough ER to the Golgi apparatus.

A somewhat different conclusion, that deletion of the C-terminal hydrophobic region alters the composition of the oligosaccharide side chains on the HA molecule, was reached from analysis of one of the Udorn truncation mutants (Sveda *et al.*, 1982). However, this mutant had lost not only the hydrophobic region but also another 110 of the 221 amino acids of HA2, including a glycosylation site, the disulphide bond which links HA1 and HA2 domains and the long alpha helix which stabilizes the structure of HA2 and forms the contacts with the other subunits

in the HA trimer. Thus, the biosynthesis and glycosylation of this mutant HA may actually have been affected by gross alterations in the structure of the molecule.

The results obtained with the anchor-minus mutants lead to the following conclusions:

(i) the hydrophobic amino acid sequence at the C-terminus of HA is required to anchor the protein in the outer membrane of the cell. Removal of this sequence converts HA from an integral membrane protein into a secretory protein. It therefore follows that signal sequences apart, there is no unique domain of amino acids common to all secretory proteins that automatically confers upon them the ability to be secreted.

(ii) the slow rate of glycosylation of the anchor-minus HA conforms to a general pattern that secretory proteins are processed more slowly than integral membrane proteins. This difference may lie at the level of transport to the Golgi apparatus or passage through it.

(iii) the signals govering the pattern of glycosylation of HA cannot lie in the 38 amino acids that normally comprise the C-terminus of the molecule.

C. Future Prospects

The experiments described above testify to the power of using *in vitro* mutagenesis of cloned HA genes to study the function of the hydrophobic domains of the protein. Many other possibilities exist to extend this technology to other portions of the molecule, including the fusion peptide, the antigenic sites, the receptor binding site and the carbohydrate attachment points. Precise dissection of the role of individual amino acids in the structure and function of the protein can be achieved by introducing single base changes at defined sites within the HA gene using oligonucleotide-directed mutagenesis (Zoller and Smith, 1983). Although such experiments will be of particular relevance to our analysis of the HA molecule, the results will in addition be of great value in understanding the structure and function of integral membrane proteins in general. Apart from the specialized features associated with its antigenicity and biological roles, the structure of the HA molecule is characteristic of a major class of cellular membrane proteins. Furthermore, because the biosynthesis of HA involves host cell enzymes and processes during translation, membrane transport, glycosylation and maturation, HA provides a profitable model for the study of membrane proteins and organelles.

Finally, with the availability of cloned DNA copies of several different membrane and secretory proteins, it is now feasible to construct genes coding for chimeric proteins, in which the various domains of one protein are exchanged for those of another. In this way it should be possible to analyse, for example, whether signal and/or anchor sequences are functionally interchangeable between different membrane proteins and whether the sequences which direct different glycoproteins to the various membranes of eukaryotic cells reside in the cytoplasmic tail, the hydrophobic anchor or the external spike of glycoprotein molecules.

References

[1] Davis, A. R., Hiti, A. L., Nayak, D. P.: Construction and characterization of a bacterial clone containing the haemagglutinin gene of the WSN strain (H0N1) of influenza virus. Gene *10,* 205–218 (1980).

[2] Davis, A. R., Nayak, D. P., Ueda, M., Hiti, A. L., Dowbenko, D., Kleid, D. G.: Expression of antigenic determinants of the haemagglutinin gene of a human influenza virus in *Escherichia coli.* Proc. Natl. Acad. Sci. *78,* 5376–5381 (1981).

[3] Emtage, J. S., Tacon, W. C. A., Catlin, G. H., Jenkins, B., Porter, A. G., Carey, N. H.: Influenza antigenic determinants are expressed from haemagglutinin gene clones in *Escherichia coli.* Nature *283,* 171–174 (1980).

[4] Gething, M. J., Bye, J., Skehel, J. J., Waterfield, M. D.: Cloning and DNA sequence of double-stranded copies of haemagglutinin genes from H2 and H3 strains elucidates antigenic shift and drift in human influenza virus. Nature *287,* 301–306 (1980).

[5] Gething, M. J., Sambrook, J.: Cell surface expression of influenza haemagglutinin from a cloned DNA copy of the RNA gene. Nature *293,* 620–625 (1981).

[6] Gething, M. J., Sambrook, J.: The expression of the influenza virus haemagglutinin gene in SV40-HA recombinants. In: Eukaryotic Viral Vectors (Gluzman, Y., ed.), 29–33. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory 1982 a.

[7] Gething, M. J., Sambrook, J.: Construction of influenza haemagglutinin genes that code for intracellular and secreted forms of the protein. Nature *300,* 598–603 (1982 b).

[8] Gluzman, Y.: SV40 transformed simian cells support the replication of early SV40 mutants. Cell *23,* 175–182 (1981).

[9] Gluzman, Y. (ed.): Eukaryotic Viral Vectors. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory 1982.

[10] Hamer, D. H., Leder, P.: Splicing and formation of stable RNA. Cell *18,* 1299–1302 (1979).

[11] Hartman, J. R., Nayak, D. P., Fareed, G. C.: Human influenza virus haemagglutinin is expressed in monkey cells using simian virus 40 vectors. Proc. Natl. Acad. Sci. *79,* 233–237 (1982).

[12] Heiland, I., Gething, M. J.: Cloned copy of the haemagglutinin gene codes for human influenza antigenic determinants in *E. coli.* Nature *292,* 851–852 (1981).

[13] Hubbard, S. C., Robbins, R. W.: Synthesis and processing of protein linked oligosaccharides *in vivo.* J. Biol. Chem. *254,* 4568–4576 (1979).

[14] Lai, C. J., Markoff, L. J., Zimmerman, S., Cohen, B., Berndt, J. A., Channock, R. M.: Cloning DNA sequences from influenza viral segments. Proc. Natl. Acad. Sci. *77,* 210–214 (1980).

[15] Lamb, R. A., Lai, C. J., Choppin, P. W.: Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. Proc. Natl. Acad. Sci. *78,* 4170–4174 (1981).

[16] Lamb, R. A., Lai, C. J.: Spliced and unspliced messenger RNAs synthesized from cloned influenza virus M DNA in an SV40 Vector: Expression of the influenza virus membrane protein (M_1). Virology *123,* 237–256 (1982).

[17] Maniatis, T., Fritsch, E. F., Sambrook, J.: Molecular Cloning–A Laboratory Manual. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory 1982.

[18] McCutchan, J. H., Pagano, J. S.: Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethyl aminoethyl-dextran. J. Natl. Cancer Inst. *41,* 351–357 (1968).

[19] Mellon, P., Parker, V., Gluzman, Y., Maniatis, T.: Identification of DNA sequences required for transcription of the human β-globin gene in a new SV40 host-vector system. Cell *27,* 279–288 (1981).

[20] Mulligan, R. C., Howard, H., Berg, P.: Synthesis of a variant β-globin in cultured monkey cells following infection with a SV40 β-globin recombinant genome. Nature *277,* 108–114 (1979).

[21] Neuberger, A., Gottschalk, A., Marshall, R. O., Spiro, R. G.: The Glycoproteins: Their Composition, Structure and Functions, 450–490. Amsterdam: Elsevier 1972.

[22] Pipas, J. M., Adler, S. P., Peden, K. W. C., Nathans, D.: Deletion mutants of SV40 that affect the structure of viral tumor antigens. Cold Spring Harbor Symp. Quant. Biol. *44,* 285–291 (1979).

[23] Perucho, M., Hanahan, D., Lipsich, L., Wigler, M.: Isolation of the chicken thymidine kinase gene by plasmid rescue. Nature *285,* 207–210 (1980).

[24] Sveda, M. M., Lai, C. J.: Functional expression in primate cells of cloned DNA coding for the haemagglutinin surface glycoprotein of influenza virus. Proc. Natl. Acad. Sci. U.S.A. *78,* 5488–5292 (1981).
[25] Sveda, M. M., Markoff, L. J., Lai, C. J.: Cell surface expression of the influenza virus haemagglutinin requires the hydrophobic carboxy-terminal sequences. Cell *30,* 649–656 (1982).
[26] Tabas, I., Kornfeld, I. J.: Purification and characterization of a rat liver golgi-mannosidase capable of processing aparagine-linked oligosaccharides. J. Biol. Chem. *254,* 11655–11663 (1979).
[27] Tooze, J. (ed.): DNA Tumor Viruses, 2nd ed. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory 1980.
[28] White, J. M., Helenius, A. A., Gething, M. J.: The haemagglutinin of influenza virus expressed from a cloned gene promotes membrane fusion. Nature *300,* 658–659 (1982).
[29] Wiley, D. C., Skehel, J. J., Waterfield, M. D.: Evidence from studies with a cross-linking reagent that the haemagglutinin of influenza virus is a trimer. Virology *79,* 446–448 (1977).
[30] Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Urlaub, G., Chasin, L.: DNA-mediated transfer of the adenine phosphoribosyl-transferase locus into mammalian cells. Proc. Natl. Acad. Sci. *76,* 1373–1376 (1979).
[31] Young, J. F., Desselberger, U., Graves, P., Palese, P., Shatzman, A., Rosenberg, M.: Cloning and expression of influenza virus genes. In: The Origin of Pandemic Influenza Viruses (Laver, W. G., ed.). New York: Elsevier/North-Holland 1983 (in press).
[32] Zoller, M. J., Smith, M.: Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 derived vectors. Methods in Enzymology, in press (1983).

7

Mutants of Influenza Virus

B. W. J. Mahy

How puzzling all these changes are!

Lewis Carroll,
Alice in Wonderland, Chapter 5

I. Introduction

Influenza virus was the first animal virus with which genetic recombination was demonstrated, when Burnet and Lind (1949), taking advantage of naturally occurring variation, found recombination in mouse brain between the neurotropic strain NWS and the non-neurotropic strain WSM. Further quantitative studies by Burnet and co-workers in Australia and by Hirst and Gotlieb in the U.S.A. established that recombination between any two strains of influenza virus occurred at a high frequency, around 35—40% (Burnet and Lind, 1952; Hirst and Gotlieb, 1955). The reason for this high frequency recombination (gene reassortment) was not suggested until 1961, when Barry (1961) demonstrated the phenomenon of multiplicity reactivation of UV-inactivated influenza virus and proposed that the virus genome might exist as "six independent radiation sensitive units". Hirst (1962) reached a similar conclusion on the basis of experiments in which the plaque-forming capacity of UV-inactivated WSN strain influenza virus was rescued by a variety of non-plaque-forming strains (Simpson and Hirst, 1961). Chemical inactivation studies of the influenza virus genome also supported the notion that there were several independent genetic units (Scholtissek and Rott, 1964). Physical analyses of the influenza virus genome carried out at this time were consistent with the idea that it was fragmented, since RNA species larger than about 18s-20s could not be obtained from virions (Davies and Barry, 1966; Duesberg and Robinson, 1967; Pons, 1970), but ten years were to elapse before the true structure of the influenza virus genome, consisting of eight single stranded RNA pieces, was established (reviewed by Palese, 1977; Barry and Mahy, 1979; Chapter 2, this volume).

In the meantime, Fenner and Sambrook (1964), reviewing the field of animal virus genetics, had stressed the importance of conditional lethal mutants in studies of the physiology of virus replication, and in the same year the first isolation of temperature-sensitive *(ts)* mutants of an animal virus (Sindbis) was reported (Burge

and Pfefferkorn, 1964). The success of these studies and the availability of a plaque-forming virus strain (WSN), led Simpson and Hirst (1968) to the isolation and characterization of the first *ts* mutants of influenza virus, which provided the basis for all further work in this field. They demonstrated high frequency recombination between pairwise crosses of certain *ts* mutants, which could be arranged into five groups. They also pointed out that it should be possible to establish the number of genome RNA pieces using this approach and to correlate certain functions with particular genome fragments by biochemical studies. The extent to which these objectives have been realised is the subject of this review; the earlier genetic studies have been well summarized by Kilbourne (1963), Hoyle (1968), and Suguira (1975), and will not be reconsidered here.

II. Characterization of Virus Mutants

Mutations in the virus genome are normally recognized by an alteration in phenotype with respect to "wild-type" virus. The phenotypic changes may in turn affect a general property of the virus, such as plaque size and morphology (growth rate), drug resistance, host range, or pathogenicity. However, the use of such markers is necessarily of limited experimental value, since they are confined to the one or two genes that affect the particular phenotypic property. The most useful class of mutants for studying the physiology of replication are conditional-lethal mutants, in which the defective virus protein functions normally under optimal conditions but is unable to function under non-permissive conditions. Theoretically, such mutations can occur in all virus genes, which makes these mutants especially valuable for mapping studies. Conditional lethal mutants which fail to replicate at a defined supra-optimal temperature (*ts* mutants) have been of enormous value in genetic studies with animal viruses. Another type of conditional lethal mutant, so-called suppressor-sensitive mutants, has been widely used in bacteriophage studies (Hartman and Roth, 1973; Glass *et al.*, 1982). These mutants fail to grow in most host cell types (which support the growth of wild-type virus) because of a nonsense mutation which terminates synthesis of an essential polypeptide. However, a few cell strains are permissive for mutant virus growth as they contain suppressor tRNAs which allow read-through at the nonsense mutation. As yet, no suppressor tRNAs have been found in animal cells. Some animal virus mutants with a suppressor-sensitive termination mutation have been recognized and studied *in vitro* using yeast suppressor tRNAs (Cremer *et al.*, 1979).

A prerequisite for genetic studies is the careful selection of a suitable virus-host cell system, and the isolation of mutants from a single cloned stock of wild-type virus (Fenner and Sambrook, 1964). For influenza virus, these requirements are not easy to satisfy, since (a) although many different cell types will act as hosts for influenza virus, most virus-cell combinations give rise to an abortive replication cycle so that the virus does not form plaques, and (b) the high rates of mutability and genetic variation of the influenza virus genome pose particular problems in defining the nature of the wild-type virus population.

A. Influenza Virus-Cell Systems

Meaningful genetic studies require the ability to plaque the virus on the host cell used. This has largely limited influenza virus *ts* mutant studies to the WSN virus strain in infections of chick embryo fibroblast cells (Simpson and Hirst, 1968; Hirst, 1973; Mackenzie, 1970 a, b) or MDBK cells (Sugiura *et al.*, 1972, 1975); to an NWS-A1/CAM recombinant, X-3311, which plaques in a human conjunctival cell line, 1-5C-4 (Ueda, 1972); to avian influenza A (fowl plague) virus infection of chick embryo fibroplast cells (Markushin and Ghendon, 1973; Scholtissek and Bowles, 1975; Almond *et al.*, 1979); or to some human candidate vaccine strains grown in primary chick, monkey or bovine kidney cells (Maassab, 1967; Mills and Chanock, 1971). More recently, MDCK cells have also been used as the host for growing human vaccine strains (Kendal *et al.*, 1979; Shimizu *et al.*, 1982, 1983).

Since it is now clear that the failure of many influenza virus strains to form plaques on a particular host cell monolayer may be overcome by the addition of trypsin (Appleyard and Maber, 1974; Klenk *et al.*, 1975), it should be possible in the future to extend the range of virus-cell systems used for genetic analyses.

B. Nature of the Virus Population

Any preparation of influenza virus is made up of a mixture of infectious and non-infectious particles, the total number of which can conveniently be scored by a haemagglutination assay or by a particle count in the electron microscope. The original estimates (Donald and Isaacs, 1954) suggested that 10 virus particles were equivalent to one infectious dose as measured by inoculation into the fertile egg, the most efficient assay system available.

A recent reappraisal of particle counts using fowl plague virus (Ghendon *et al.*, 1983) yielded values (10 particles per EID_{50}) almost identical to the original determinations by Isaacs and Donald (1955). Under ideal conditions, preparations of fowl plague virus, for example, can be obtained where one chick embryo fibroblast cell plaque-forming unit is equivalent to one EID_{50} (Carter and Mahy, 1982 a), but usually plaque-forming ability is less efficient than infectivity in eggs. There is, however, considerable evidence that the non-plaque-forming particles interact genetically, and in certain circumstances may complement each other. The non-plaque-forming particles presumably consist of physically damaged (*e.g.* thermally inactivated) particles as well as particles which lack a complete set of genome RNA segments. Provided more than one such particle enters a cell, complementation may occur; this phenomenon was demonstrated experimentally with UV-inactivated virus many years ago and referred to as multiplicity reactivation (Henle and Liu, 1951; Barry, 1961). Aggregation or clumping of virus particles increases the possibility of complementation between non-infectious particles (Hirst and Pons, 1973; Dales and Pons, 1976). Biological activity of such non-infectious particles was also detected in studies of von Magnus incomplete influenza virus, revealing a sub-population of defective virions which do not replicate or interfere with the growth of

standard virus but are capable of mRNA transcription (defective non-interfering particles) (Carter and Mahy, 1982 a).

Defective influenza virions which can interfere with the growth of standard virus (von Magnus, 1954; see chapter 8, this volume) also contribute to most influenza virus populations, since they are readily perpetuated and amplified as a consequence of multiple particle infection. A virus population containing such particles is said to be incomplete; a virus population without such defective virions is known as standard virus. Incomplete virions are characterized by alterations in the genome structure, which includes loss of, or internal deletions in, the three largest RNA segments (Nayak, 1980; Pons, 1980; Carter and Mahy, 1982 b) and recent evidence suggests that deletions and sequence rearrangements can be generated within any RNA segment (P. Jennings, J. T. Finch, G. P. Winter and J. S. Robertson, submitted). High multiplicity infection of cells in tissue culture, even with standard virus containing no detectable incomplete virions, generates a specific reduction in RNA segments 1–3 during the first multiplication cycle (Carter and Mahy, 1982 b). The mechanism by which such aberrant forms are generated is not yet clear. The possible interaction of all the components in a virus population must therefore be considered when analyzing defects in influenza virus mutants, particularly if high multiplicities of infection are employed.

C. Natural Mutation Rate

Recent advances in genetic manipulation and nucleotide sequencing techniques have permitted more detailed characterization of the mutation rates in RNA virus genomes than had hitherto been feasible. Such studies indicate that RNA virus genomes have mutation rates which can be a million-fold higher than the rate of mutation found in the host DNA chromosome (Holland *et al.*, 1982). For this reason, it is relatively easy to isolate spontaneous mutants from RNA virus populations. Temperature-sensitive mutants of influenza virus arising spontaneously during laboratory passages were originally used by Simpson and Hirst (1968) and others were isolated in later studies (Hirst, 1973; Almond *et al.*, 1979). There are also several reports describing the isolation of naturally-occurring *ts* mutants from the human population (Oxford *et al.*, 1980; Chu *et al.*, 1982; Zhang *et al.*, 1982). However, a *ts* phenotype may occur with only a small proportion of the total number of mutants in the virus population, and the true natural mutation rate may be much higher. Brand and Palese (1980) identified mutants of influenza A virus by oligonucleotide mapping; after 12 plaque to plaque passages all eight virus clones studied contained base changes in their RNA. However this technique cannot measure the precise error rate, as only about 10% of the genome is analyzed in T_1 ribonuclease mapping (Palese and Young, 1982).

As an alternative approach to measure mutation rates, Portner *et al.* (1980) used monoclonal antibodies and compared the rates of mutation to antibody resistance of influenza virus, vesicular stomatitis virus, and Sendai virus. All three viruses gave mutation rates of about $10^{-4.5}$ per replication, whereas the error rate of E. coli DNA replication has been estimated as 10^{-8} to 10^{-10} (Fowler *et al.*, 1974).

Recently, the genetic variability of PR8 strain influenza virus RNA was explored by analysis of over 200 clones prepared in derivatives of bacteriophage M13 (Fields and Winter, 1981 a). Sequence heterogeneity in the cloned DNA was present at one nucleotide difference per 3700 nucleotides (3×10^{-4}), but the base mismatch errors introduced by the DNA replication enzymes used to prepare the cDNA are probably so high as to obscure the natural variation in the influenza virus RNA population (Gopinathan *et al.*, 1979; McGeoch, 1981).

D. Induction of Mutants

The majority of *ts* mutants which have been isolated were derived by deliberate mutagenesis (Simpson and Hirst, 1968; Fields and Joklik, 1969). This procedure may result in a number of mutations for each influenza virus mutant. The likelihood of multiple mutations is further increased by plaque to plaque passages carried out whilst preparing stocks of mutants (Spindler *et al.*, 1982), and the selection of mutants with low plaquing efficiency at the non-permissive temperature will also favour selection of multiple mutants (Simpson and Hirst, 1968). For these reasons, the probability is high that "silent" mutations exist in addition to the *ts* lesion under study and this must be borne in mind when investigating a particular *ts* phenotype. Several examples where such "silent" mutations have become recognized under particular experimental conditions have been documented (Kendal *et al.*, 1979; Murphy *et al.*, 1980; Tolpin *et al.*, 1981; Thierry and Spring, 1981).

1. Mutagens

A variety of mutagens have been used to derive influenza virus mutants (see Table 1). Of these, growth of the virus in the presence of 5-fluorouracil (5-FU) has been most widely employed. Incorporation of 5-FU in place of U is efficient (up to 50% base substitution), and other bases are unchanged (Gordon and Staehelin, 1959). Replication of 5-FU-containing RNA results in both A → G and U → C transitions, because the increased ability of 5-FU to tautomerise to the enol form results in base-pairing with G instead of A (Freese, 1959).

Direct treatment of wild-type virus stocks has been carried out with nitrous acid, which results in deamination, so that C is converted to U and A to hypoxanthine (Schuster and Schramm, 1958), or with hydroxylamine, which converts C to a derivative which pairs with A (Freese *et al.*, 1961). Growth of virus in the presence of N-methyl-N-nitrosourea (NTU) or treatment of virus with N-methyl-N′-nitro-N-nitrosoguanidine (NTG) have also been used in influenza virus mutant production. These alkylating agents act principally upon guanine, and it is likely that 0-6-methylation of guanine is responsible for their mutagenic action (Loveless, 1969), which results both in transitions and transversions during nucleic acid replication. Other mutagens used recently in isolating influenza virus mutants (Shimizu *et al.*, 1982 a) are an acridine derivative, ICR 191, which has been shown to induce frame-shifts in DNA by intercalation (Ames and Whitfield, 1966), and UV-irradiation, which causes in DNA a variety of transitions and transversions and can introduce deletions (Schwartz and Beckwith, 1969).

E. Leak and Reversion

Leak and reversion are major problems in working with *ts* mutants and can seriously affect the interpretation of phenotypic analyses. Normally, *ts* mutants are selected since they grow at the permissive temperature (which varies in different sets of influenza virus mutants between 31 °C and 36 °C) but are severely restricted in growth at non-permissive temperatures (38 °C—42 °C). The molecular basis for temperature sensitivity is assumed to be a single amino-acid (mis-sense) replacement in a protein which affects its folding or stability so that it is inactive at the non-permissive temperature. By definition, *ts* mutants will only identify functions essential for virus multiplication, which may not include all protein functions. Leakiness results from the residual activity of a *ts* protein at the non-permissive temperature. The leak yield may be determined for any *ts* mutant as the ratio of plaque forming units/ml produced at the non-permissive compared to the permissive temperatures. Virus yielded at the non-permissive temperature is, of course, still temperature sensitive. Strict temperature control can minimize leak, and this can be accomplished in biochemical experiments with *ts* mutants by immersion of cell culture vessels in water baths to maintain the non-permissive temperature (Sugiura *et al.*, 1975; Ghendon *et al.*, 1973; Wolstenholme *et al.*, 1980). We have found in experiments with fowl plague virus (Rostock) *ts* mutants that gassed air incubators do not provide adequate temperature control. Clearly, the leak rate of individual *ts* mutants must be taken into account, and only those which are most stringently temperature-sensitive ("tight" mutants) should be chosen for biochemical experimentation. Leak is also a problem in genetic characterization if mutants are analyzed by complementation.

Reversion contributes to the yield from infection at the non-permissive temperature, but in this case the progeny are phenotypically wild-type. In addition to a back mutation at the original site (reverse mutation), a mutation at another site in the genome may occasionally result in loss of the *ts* phenotype (suppressor mutation). The suppressor mutation may occur on the same or a different gene from that carrying the *ts* lesion. The reversion rate varies, but usual frequencies for influenza virus mutants are 10^{-5} to 10^{-4}; this is sufficiently high that caution should be exercised in passaging mutant stocks without frequent checks to ensure maintenance of the *ts* phenotype. Scholtissek and Spring (1981 b) made a study of the reversion rates of *ts* mutants with lesions in different RNA segments. They found that the reversion rates for segments 1, 2, 3 and 5 (genes coding for NP and the three P proteins) were relatively high compared to the rates for segments 4 and 6, encoding the HA and NA glycoproteins respectively. Comparable studies have not been reported for other groups of *ts* mutants.

Almond (1978) isolated fowl plague virus mutants by the plaque enlargement method either spontaneously (Sp series), by growth in 5-FU (mF series), by treatment of wild-type virus with 100 μg/ml NTG (mN series) or with 50 μg/ml NTG (no prefix). He also isolated 5-FU-induced mutants by picking and testing every plaque (US series). His results with a total of 49 mutants suggested that:

(a) the stability of mutants was not related to the method of their isolation,

Table 1. *Temperature-sensitive influenza virus mutants*

Reference	Wild-type strain	Cells used for isolation experiments*	Temperature (°C) permissive/ non-permissive	No. of mutants	Mutagen	No. of complementation or recombination groups
Simpson and Hirst (1968) and Hirst (1973)	WSN (H1N1)	CEF/CEF	34/39.5	56	F-FU (23) NTG (9) HNO_2 (3) Spontaneous (21)	8R
Mackenzie (1970)	WSN (H1N1)	CEF/CEF or AOS	33/39	16	5-FU (11) HA (5)	5R
Sugiura *et al.* (1972 and 1975)	WSN (H1N1)	MDBK/MDBK	34/39.5	34	5-FU (31) HA (3)	7R
Mills and Chanock (1971)	A2 (H2N2)	BK/BK	32/38	2	5-FU (2)	–
Ueda (1972)	Recombinant X-3311	1-5C-4/1-5C/4	31/38	6	5-FU (6)	–
Markushin and Ghendon (1973) Ghenkina and Ghendon (1979) Ghendon and Markushin (1980)	FPV Weybridge (H7N7)	CEF/CEF	36/42	15	HA (8) 5-FU (4) HNO_2 (1) NTU (1)	5C
Scholtissek *et al.* (1974) Scholtissek and Bowles (1975) Koennecke *et al.* (1981)	FPV Rostock (H7N1)	CEF/CEF	33/40	30	5-FU (30)	7R
Spring *et al.* (1975)	Hong Kong (H3N2)	CEF/BK/RMK	33/39	9	5-FU (9)	7C
Almond (1978) Almond *et al.* (1979)	FPV Rostock (H7N1)	CEF/CEF	34/40.5	49	5-FU (7) NTG (36) Spontaneous (6)	6R
Almond (1978) Almond *et al.* (1977) Almond and Barry (1979)	FPV Dobson (H7N7)	CEF/CEF	34/40.5	5	Spontaneous (5)	–
Shimizu *et al.* (1982 a, b)	A/Udorn/72 (H3N2)	BK/RMK	34/40	83	ICR191 HNO_2 UV	13C 8R

(b) stable, non-leaky mutants were obtained by the plaque enlargement method,

(c) two out of three US mutants tested were quite leaky,

(d) reversion rates among the mutants varied independently of leakiness.

Clearly there are no hard and fast rules for obtaining tight mutants with low reversion rates, but ideally each mutant should be subjected to careful analysis before putting it to experimental use.

III. Temperature-Sensitive Mutants

The principal aim of isolating and studying *ts* mutants is to investigate the function and mode of replication of the virus genome. The main groups of influenza virus *ts* mutants are listed in Table 1. Since several groups have derived mutants from the same virus strain, some confusion has developed regarding mutant nomenclature. It would seem desirable to introduce some identification system so that, for example, fowl plague virus Rostock *ts 3* isolated by the Giessen group can be readily distinguished from *ts 3* of the same virus strain isolated in Cambridge. Table 2 describes a proposed nomenclature which I shall adopt in this review, and which may find wider acceptance.

A. Genetic Interaction and Classification of Mutants

The use of pair-wise crosses for genetic analysis of *ts* mutants as introduced by Simpson and Hirst (1968) provides a reliable method of grouping which fully exploits the segmented nature of the genome. With one possible exception (Scholtissek *et al.*, 1978 b; Rohde and Scholtissek, 1980) a true recombination event

Table 2. *Designation of virus mutants by laboratory of origin*

Strain	Investigators	Place	Code letters	Example
Fowl plague virus	Markushin and Ghendon (1973)	Moscow	M	*ts* M43
	Scholtissek and Bowles (1975)	Giessen	G	*ts* G3
	Almond *et al.* (1979)	Cambridge	C	*ts* C3
WSN	Simpson and Hirst (1968)	New York	NY	*ts* NY3
	Mackenzie (1970)	Canberra	CAN	*ts* CAN3
	Sugiura *et al.* (1972 and 1975)	New York and Tokyo	NYT	*ts* NYT3
Hong Kong	Spring *et al.* (1975 b)	Bethesda	B	*ts* B315
Udorn	Shimizu *et al.* (1982 a, b)	Bethesda	BU	*ts* BU1299

within an RNA segment has not yet been demonstrated for influenza virus, although it has been shown to occur in the single-stranded RNA genome of a picornavirus (King *et al.*, 1982). Free reassortment of all segments during a dual infection between two *ts* mutants at the non-permissive temperature thus gives high recombination (strictly, reassortment) frequencies if the mutations are on different segments and insignificant levels of recombination if the lesions are on the same segment. The recombination frequency is derived by doubly infecting cells at the permissive temperature with the *ts* parents (a and b) and plating out the yield at both permissive (ab^P) and non-permissive (ab^{NP}) temperatures. To check for reversion, single infections at the non-permissive temperature should be carried out and those yields recorded (a^{NP} and b^{NP}). The recombination frequency (per cent ts^+ [wild-type] progeny) may then be calculated as:

$$\% \; ts^+ = \frac{ab^{NP} - a^{NP} - b^{NP}}{ab^P} \times 100$$

Several groups of investigators, using different virus strains, have confirmed that this recombination event yields stable wild-type progeny from which the *ts* phenotype has been eliminated (Simpson and Hirst, 1968; Sugiura *et al.*, 1975; Markushin and Ghendon, 1973; Almond, 1978).

Simpson and Hirst (1968) also demonstrated complementation between certain pairs of *ts* mutants, where mixed infections carried out at the non-permissive temperature gave yields more than 100-fold greater than single infections. In this case, the progeny virus consists of a mixture of *ts* viruses with ts^+ recombinants. Although pairs of *ts* mutants may be screened more rapidly by complementation rather than recombination testing, there is now clear evidence that complementation may occur between mutants having lesions within the same RNA segment, or even within the same gene. Such intragenic complementation has been reported for FPV Rostock mutants *ts* G18 and *ts* G236 which have a *ts* lesion in RNA segment 2 (Heller and Scholtissek, 1980), with two recombinant H3N2 viruses also carrying RNA segment 2 lesions (Massicot *et al.*, 1982), with segment 5 and 6 in a set of A(Hong Kong)68 mutants (Thierry *et al.*, 1980) and with segments 2, 3, 6, and 8 in studies of a series of A/Udorn/72 (H3N2) mutants (Shimizu *et al.*, 1981).

An additional problem with complementation analysis arises from suppression of the *ts* phenotype by a different segment, presumably coding for a protein which complements the functional defect in the segment carrying the *ts* lesion. Such extragenic suppression was first noted in reovirus *ts* mutants (Ramig and Fields, 1979), and several examples have now been recorded with influenza virus mutants. Tolpin *et al.* (1981) found that a mutation in the PA gene of a recombinant human influenza strain suppressed the expression of a *ts* phenotype of the PB2 gene. This is an example of a "suppressor mutation" analogous to the result of Ramig and Fields (1979). However, in recombination experiments with a fowl plague virus/virus N recombinant (19N) carrying a *ts* lesion in RNA segment 8, Scholtissek and Spring (1981 b, 1982) found that ts^+ recombinants formed with other influenza strains had frequently replaced segments 1 or 2 but not segment 8. This suggests that gene product(s) of segments 8 (non-structural proteins) must cooperate with products of segments 1 and 2 (polymerase proteins), during replication. Such recombinants

were termed "suppressor recombinants" by Scholtissek and Spring (1981 a, b). A similar extragenic suppression was revealed during comparison of two groups of fowl plague virus mutants, isolated in Moscow (Weybridge strain) and in Cambridge (Rostock strain). When mutant *ts* M29 was crossed with *ts* C44, both mutants having a *ts* lesion in RNA segment 1, ts^+ recombinants appeared at high frequency. These ts^+ recombinants retained the gene carrying a *ts* mutation, and further analysis showed that segment 2 of the Weybridge strain had suppressed the effect of the *ts* lesion on segment 1 of the Rostock strain (Ghendon *et al.*, 1982 b).

The existence of intragenic suppression will clearly influence the results obtained from complementation tests of groups of mutants, although recombinant analysis should not be so affected. However, during analysis of *ts* mutants by recombination, the possibility of extragenic suppression must be borne in mind, as this phenomenon may seriously hinder the interpretation of results (Fields, 1981; Ghendon *et al.*, 1982 b).

Two other reports of factors which may affect genetic analyses are the independent segregation of *ts* growth from *ts* polymerase activity found during study of the A/Ann Arbor/6/60 cold mutant (Kendal *et al.*, 1979), and evidence for a mutation which enhances the mutation rate of the vesicular stomatitis virus genome (Pringle *et al.*, 1981). The latter phenomenon has not yet been observed with influenza viruses, however.

B. Assignment of *ts* Lesions to Individual Genome Segments

In 1976 a number of laboratories independently demonstrated, using high resolution polyacrylamide gel electrophoresis in the presence of urea, that the influenza virus genome could be fractionated into eight species (Bean and Simpson, 1976; McGeoch *et al.*, 1976; Palese and Schulman, 1976 a; Pons, 1976; Ritchey *et al.*, 1976 a; Scholtissek *et al.*, 1976). These species were shown to be distinct, for fowl plague virus, by T_1 oligonucleotide fingerprinting (McGeoch *et al.*, 1976) and later by direct terminal sequence analysis (Robertson, 1979). The close correlation between the number of genome segments and the number of high frequency recombination groups observed in *ts* mutant studies was immediately obvious, and several laboratories developed methods of assigning *ts* mutations to individual genome RNA segments. Most of these methods exploited the fact that different influenza virus strains may show migrational differences in virus-specific RNA's and polypeptides when subjected to electrophoresis in polyacrylamide gels (Palese and Schulman, 1976 a; Ritchey *et al.*, 1977; Almond *et al.*, 1977; Bean and Webster, 1978).

1. Cambridge Mutants

Almond *et al.* (1977) used *ts* mutants of two strains of fowl plague virus, Rostock (H7N1) and Dobson (H7N7), to generate ts^+ recombinants by mixed infection at the permissive temperature. Recombinants detected by plaquing the yield at the non-permissive temperature were then analyzed by polyacrylamide gel electrophoresis to determine the parental origin of RNA segments and intracellular induced polypeptides. Assuming that the Dobson and Rostock parent viruses have

ts lesions in different segments, ts^+ progeny containing the Dobson RNA segment which corresponds to the defective Rostock gene, and *vice-versa,* will arise at high frequency. But none of the other six segments is subject to selection, so the procedure generates 2^6 or 64 RNA composition classes of ts^+ progeny (Almond *et al.*, 1977). To obviate the analysis of very large numbers of recombinants, one parent (in this case the *ts* Dobson virus) was irradiated with UV light to an extent estimated to leave one genetically competent RNA segment per particle. This resulted in ts^+ progeny which usually had replaced only the *ts* segment of Rostock virus with a Dobson segment (Almond *et al.*, 1977; Almond and Barry, 1979). It was possible by this technique to map all six recombination groups of fowl plague virus Rostock *ts* mutants to individual genome RNA segments (Table 3).

2. *New York-Tokyo Mutants*

Palese *et al.* (1977 b) also developed a method for *ts* lesion assignment, which exploited gel migrational differences between the RNA segments of the PR8 and Hong Kong strains of influenza virus (Palese and Schulman, 1976 a, b; Ritchey *et al.*, 1976 a, b; Palese, 1977). In applying this method to *ts* mutants of the WSN strain (H1N1) (isolated by Sugiura *et al.*, 1972, 1975), Palese *et al.* (1977 b) generated ts^+ recombinants by mixed infection of MDCK cell monolayers with a WSN *ts* mutant and ts^+ Hong Kong virus at the non-permissive temperature in the presence of trypsin. Recombinant yields were then analyzed by plaque assay at non-permissive temperature in MDBK cells, in which the Hong Kong virus does not replicate (Schulman and Palese, 1976). By this technique, complementation-recombination groups of five of the Sugiura *ts* mutants (I, II, III, V, and VII) were assigned to RNA segments (Table 3) (Palese *et al.*, 1977 b; Ritchey and Palese, 1977). Previous studies of the functional defects in complementation-recombination groups IV and VI enabled them to be mapped onto segments 6 (neuraminidase) and 4 (haemagglutinin) respectively (Palese *et al.*, 1974; Ueda and Kilbourne, 1976; Palese, 1977). A mutant (*ts* NYT 60S) apparently belonging to an eighth complementation-recombination group was described by Nakajima and Sugiura (1977), but further analysis has shown that this assignment was incorrect (A. Sugiura, personal communication).

3. *Giessen Mutants*

For analysis of fowl plague virus Rostock (H7N1) *ts* mutants isolated in Giessen, recombinants were prepared by mixed infection of chick embryo fibroblast cells at relatively high multiplicity with a single *ts* mutant (40 PFU per cell) and a human influenza ts^+ strain (10 PFU per cell). Advantage was taken of the fact that most human strains do not plaque in chick embryo fibroblast cells, so that all plaques obtained at the non-permissive temperature (except for occasional *ts* mutant revertants) must be recombinants. These progeny were then passaged thrice more by plaquing at the non-permissive temperature. To analyze the gene composition of these recombinants, a hybridization method was used in which ^{32}P-labelled RNA segments of the parent fowl plague virus Rostock are hybridized to saturating amounts of unlabelled complementary RNA (cRNA) of the ts^+ recombinant. The

Table 3. *Results of mapping influenza* ts *mutant groups onto individual RNA segments*

Genome RNA segment	Complementation-recombination group					
	FPV (Rostock)		FPV (Weybridge)	WSN	Hong Kong	Udorn
	Cambridge[a]	Giessen[b]	Moscow[c]	New York and Tokyo[d]	Bethesda[e]	Bethesda[f]
1	III	I	C	I	1	A
2	II	III	D	III	5	B
3	V	II	B	II	6	C
4	VI	V	–	VI	–	–
5	I	VI	–	V	4	–
6	–	IV	E	IV	2, 3, 7	F
7	–	–	F	VII	–	–
8	IV	VIII	–	–	–	H

[a] Almond *et al.*, 1979; [b] Scholtissek and Bowles, 1975; [c] Markushin and Ghendon, 1973; [d]Sugiura *et al.*, 1972, 1975.
[e] Spring *et al.*, 1975; [f] Shimizu *et al.*, 1982 a, b.

cRNA is prepared by incubating a microsome preparation, purified from cells five hours post-infection with the recombinant, with all four nucleoside triphosphates and co-factors (Scholtissek and Rott, 1969 a). After a 20 minute reaction period, the microsomes are sedimented at 40,000 × g and the RNA extracted from the supernatant fluid is mainly cRNA (Scholtissek and Rott, 1970). For most genome RNA segments, the degree of hybridization between ^{32}P-labelled fowl plague virus RNA and a human influenza strain cRNA as detected by ribonuclease sensitivity is not more than 80%, compared to 100% in the homologous reaction. However, the sensitivity of the method can be greatly increased if, after hybridization and before ribonuclease treatment, the samples are heated for 10 min at 75 °C in the presence of 1% formaldehyde, to melt partial hybrid structures. The latter procedure is essential with highly conserved segments such as 7 and 8, but can introduce more variability in results than is seen without formaldehyde melting (Scholtissek *et al.*, 1976). By this method of molecular hybridization the various RNA segments in ts^+ recombinants could be assigned to one or other parent, and the segment carrying the *ts* defect identified (Table 3). A further refinement introduced by the Giessen group was to compare the proteins induced by the parent and recombinant viruses by tryptic peptide mapping, which enabled complete genetic maps of fowl plague virus Rostock and virus N to be obtained (Harms *et al.*, 1978; Rohde *et al.*, 1977, 1978).

4. Moscow Mutants

The mutants of fowl plague virus Weybridge (H7N7) originally isolated by Markushin and Ghendon (1973), have recently been assigned to RNA segments using a procedure for recombinant virus RNA analysis developed by Hay *et al.* (1977 a). Recombinants were prepared between *ts* mutants and a human strain A/Krasnodar/101/59 (H2N2) (Ghendon *et al.*, 1979 a). The recombinants were used to infect chick embryo fibroblast cells in the presence of cycloheximide, which inhibits the synthesis of vRNA (Hay *et al.*, 1977 a, b). Virus-specific cRNA was then labelled with ^{3}H-uridine from 1—4 hours post-infection in the presence of cycloheximide, and after extraction was hybridized with unlabelled virion RNA (cRNA) of either parental type. The resultant double-stranded RNAs, after treatment with S_1 nuclease, were analyzed by electrophoresis in 4% polyacrylamide gels. If vRNA segments are homologous, the double-stranded RNAs survive S_1 nuclease treatment, but if not, the partial hybrids are destroyed. On this basis, the origin of recombinant virus RNA segments could readily be determined (Ghendon and Markushin, 1980; Ghendon *et al.*, 1983 a) (Table 3). Certain of these *ts* lesion assignments have also been confirmed by recombination analysis with the Cambridge fowl plague virus Rostock mutants (Ghendon *et al.*, 1982 b).

5. Bethesda Mutants

The *ts* mutants of influenza A/Hong Kong/68 (H3N2) isolated by Spring *et al.* (1975 b) have recently been analyzed by recombination at the permissive temperature with influenza A/WSN virus. In this case, *ts* recombinants were

selected by incubating the yield with a hyperimmune anti-H3 antiserum before plaquing onto MDCK cell monolayers in the presence of trypsin (Thierry *et al.*, 1980). The genome composition of the *ts* recombinants was analyzed by exploiting gel migrational differences of the parent strains essentially as carried out by Palese and Schulman (1976 a), except that the RNA bands were visualized by ethidium bromide staining and UV illumination. By these techniques, complementation groups 4 and 3 of the Hong Kong mutants were assigned to RNA segments 5 (neuraminidase) and 6 (nucleoprotein) respectively (Thierry *et al.*, 1980). However both of these mutant groups were able to complement influenza WSN *ts* NYT mutants carrying lesions on the same genome segments, and it was concluded that these results were due to intracistronic complementation (Thierry *et al.*, 1980).

In a later study, Thierry and Spring (1981) used PR/8/34 (H1N1) as the ts^+ parent to derive a further series of *ts* recombinants with the Hong Kong *ts* mutants. After determining which segments carried *ts* defects, they assigned Hong Kong complementation group 6 to RNA segment 3, and complementation groups 2 and 7 (in addition to group 3, Thierry *et al.*, 1980) to RNA segment 6. The existence of three complementation groups carrying lesions on segment 6 encoding the nucleoprotein was taken to indicate that NP is a multifunctional polypeptide (Thierry and Spring, 1981). These results, together with previous genetic and biochemical mapping studies (Palese and Ritchey, 1977; Massicot *et al.*, 1980 a), allowed the assignment of the Hong Kong *ts* mutants to five of the RNA segments (Table 3).

Another group of Bethesda *ts* mutants, derived from the A/Udorn/72 (H3N2) strain (Shimizu *et al.*, 1982 a) have been grouped into eight recombination groups and partially assigned to individual genome RNA segments by complementation-recombination tests with WSN (NYT) mutants the lesions of which had already been mapped. To date, A/Udorn mutants have been found with lesions in segments 2, 3, 6, and 8 encoding the P1, P2, NP and NS polypeptides respectively (Shimizu *et al.*, 1982 b).

6. *Summary*

The variety of techniques employed in the assignment of *ts* mutant groups to individual RNA segments mainly reflects the technical expertise available in the various laboratories handling these mutants. Clearly, as the recent analysis of Bethesda mutants has shown, it is possible to make short cuts in the gene assignment now that well characterized WSN mutants, for example, are available. Similarly, it has been useful to compare the Moscow fowl plague virus mutants with those generated in Cambridge. However, as the Bethesda group have clearly shown (Thierry *et al.*, 1980; Shimizu *et al.*, 1982 b), there are pitfalls in comparisons between *ts* mutants, particularly where complementation is used. The various biochemical and genetic interactions which may lead to intrasegmental complementation, referred to earlier and discussed at length by Shimizu *et al.* (1982 b) must always be borne in mind when assigning *ts* lesions to particular genome RNA segments by recombination analysis.

C. Phenotypic Analysis of Temperature Sensitive Mutants

The principal aim in isolating and characterizing *ts* mutants of animal viruses is to further our understanding of the functional organization of the virus genome. With influenza virus, rapid advances in RNA and DNA sequence analysis, since the eight genome segments were first characterized in 1976, have led to determination of the complete primary structure (13, 588 nucleotides) of the genome of the A/PR/8/34 (H1N1) strain by Winter and colleagues (Winter and Fields, 1980, 1981, 1982; Fields and Winter, 1981 b, 1982; Fields *et al.*, 1981; Winter *et al.*, 1981 a, b). Genome segments of several other influenza A strains have also been sequenced; for fowl plague virus Rostock (H7N1), for example, the primary sequences of 5 of the 8 segments have now been determined (Porter *et al.*, 1979, 1980; McCauley *et al.*, 1982; J. S. Robertson and I. Roditi, in preparation), the exceptions being segments 2, 5, and 6. The use of *ts* mutants is thus given a firm structural basis from which functional interactions of the gene products may be studied.

Two sets of WSN strain mutants, those isolated in Canberra by Mackenzie (1970 a) and in New York by Simpson and Hirst (1968) and Hirst (1973), have not so far been assigned to RNA genome segments. The same is true of the X-3311 mutants isolated in Tokyo by Ueda (1972). For this reason, although some phenotypic analyses have been made with these mutants, it is difficult to correlate function with structure, and so this earlier work will not be reviewed here. An excellent account of the phenotypic characterization of these three sets of mutants was given by Hightower and Bratt (1977).

1. Assigment of Polypeptides to RNA Segments

Table 4 summarizes current information on the genome structure of the PR8 virus strain, and gives the coding assignments which have been deduced for the various strains from which sets of mutants have been derived. Although differences in electrophoretic migration of certain influenza virus-specific polypeptides, notably the P polypeptides, have been observed in different laboratories, it is apparent that the migration of virus RNA segments under denaturing conditions is much less variable. Thus sequence comparisons show that segments 1 and 3 of fowl plague virus Rostock (Robertson, 1979), correspond to segments 1 and 3 of A/PR/8/34 on the basis of limited terminal sequences (Desselberger *et al.*, 1980). The complete sequences of genes 1, 2, and 3 of the PR8 strain, which were ordered not by gel migration but by comparison with fowl plague virus terminal sequences (Winter and Fields, 1982; Fields and Winter, 1982), reveal that the gene products of segments 1, 2, and 3 are a basic protein PB2 (85,900 m.w.) a larger basic protein PB1 (86,500 m.w.) and an acidic protein PA (82,400 m.w.) respectively. This is in complete agreement with two-dimensional gel analysis of the products of those genes from fowl plague virus Rostock (Horisberger, 1980; C. R. Penn, D. Blaas, E. Kuechler, and B. W. J. Mahy, unpublished observations) from the WSN influenza virus strain (Horisberger, 1980; Ulmanen *et al.*, 1981) and from the PR8 influenza virus strain (Blaas *et al.*, 1982 a, b). In each case the basic proteins separate as a fast migrating (PB2) and slow migrating (PB1) polypeptide.

Table 4. *Influenza genome RNA segments and their coding assignments*

Segment	Length[a] (nucleotides)	mRNA length[b] (nucleotides)	Predicted nascent[a] polypeptide length (amino-acid residues)	Encoded polypeptide					
				PR8[c]	WSN[c]	FPV (Rostock)[d]	FPV (Weybridge)[e]	Hong Kong[c]	Udorn[c]
1	2341	2320	759	PB2[f]	PB2	PB2	PB2?[g]	PB2	PB2
2	2341	2320	757	PB1	PB1	PB1	PB1?	PB1	PB1
3	2233	2210	716	PA	PA	PA	PA?	PA	PA
4	1778	1756	566	HA	HA	HA	HA	HA	HA
5	1565	1539	498	NP	NP	NP	NP	NP	NP
6	1413	1391	454	NA	NA	NA	NA	NA	NA
7	1027	1004	252	M_1	M_1	M_1	M_1?	M_1	M_1
		314	97	M_2	M_2	M_2	M_2?	M_2	M_2
		275	9	?	?	?	?	?	?
8	890	868	237	NS_1	NS_1	NS_1	NS_1	NS_1	NS_1
		396	121	NS_2	NS_2	NS_2	NS_2	NS_2	NS_2

[a] Determined from sequence analysis by Winter and colleagues (see text).
[b] Deduced from the virion RNA sequence, excluding host cell primer and poly(A) tract.
[c] Based on recombinant analyses by Palese and colleagues (Palese, 1977) and Lamb and Choppin (Lamb, this volume).
[d] From Inglis *et al.* (1977); Almond *et al.* (1977, 1979); Almond and Barry (1979); Inglis *et al.* (1980); Inglis and Brown (1981).
[e] Ghendon *et al.* (1979); Ghendon *et al.* (1982 d).
[f] Based on mobilities in two-dimensional gels as shown by Horisberger (1980), Blaas *et al.* (1982 b) and Ulmanen *et al.* (1981).
[g] ? Indicates those assignments not yet certain.

The only other strain differences concern genome RNA segments 5 and 6, which code for the nucleoprotein and neuraminidase genes respectively. In some virus strains, *e.g.* fowl plague virus Dobson and the human Hong Kong strain, the migrational order of these RNAs is reversed on polyacrylamide/urea gels.

2. *Mutants with* ts *Lesions in RNA Segment 1*

RNA segment 1 encodes a basic polypeptide (PB_2) which in the PR8 virus strain has a molecular weight of 85,900 (Fields and Winter, 1982). Recent biochemical studies have established one of the functions of this protein. During the initial stages of mRNA synthesis in infected cells, a host cell mRNA molecule containing a methylated cap 1 structure (m7G[5′]ppp[5′]N_m) is recognized by the viral transcriptase complex and may be used in priming transcription. PB2 has now been shown to act as a cap-recognizing protein *in vitro* both by cross-linking with UV-light (Ulmanen *et al.*, 1981) and by photo-affinity labelling (Blaas *et al.*, 1982 a, b; Penn *et al.*, 1982). The latter technique has the advantage that it does not inactivate the RNA transcription complex. In all four virus strains studied so far, PR8, WSN, fowl plague virus Rostock, and fowl plague virus Dobson, PB2 was found to be a cap-recognizing protein.

This suggests that the gene product of RNA segment 1 functions in the early stage of virus infection, and is involved in the synthesis of messenger RNA. A variety of phenotypic studies of influenza virus *ts* mutants with lesions in genome RNA segment 1 provide support for this conclusion.

a) Virion RNA Transcriptase Activity

Analysis of the virion RNA transcriptase activities of the Cambridge *ts* mutants of fowl plague virus Rostock showed that segment 1 encoded a protein which was involved in priming the virion transcriptase reaction. Nichol *et al.* (1981) analyzed the virion transcriptase activity of eleven *ts* mutants from six different recombination groups. The assays were carried out either at 31 °C (the optimal temperature for RNA transcriptase activity) or at 40.5 °C (the non-permissive temperature). At the latter temperature, the RNA transcriptase activity of wild-type fowl plague virus, measured in the presence of the dinucleotide adenylyl (3′ and 5′) guanosine (ApG) as primer, was 10—13% of the activity measured at 31 °C. Analysis of *ts* mutants under the same conditions gave similar values at the two temperatures for all except one, *ts* C17 (group III) which gave only 1.4% activity at 40.5 °C compared to 31 °C. This mutant, carrying a *ts* lesion in the gene encoding PB2, was considered to be defective in ApG-primed *in vitro* RNA transcriptase activity. Further analysis using rabbit globin mRNA as primer in place of ApG showed that both *ts* C17 and *ts* C44 were negative for RNA transcriptase activity, whereas the other nine mutants tested (from other recombination groups) were positive. Apparently *ts* C44 (also a group III mutant) is able to transcribe virion RNA *in vitro* in the presence of ApG (complementary to the 3′ termini of the virion RNA segments), but fails to recognize (or utilize) a capped mRNA molecule as a primer. When these studies were extended to include four other group III mutants, all were found to be defective in mRNA-primed transcription *in vitro;* three (*ts* C3, *ts* C5, and *ts* C21) were also negative at 40.5 °C in ApG-primed transcription, but one (*ts* C9) displayed a

similar *ts* phenotype to *ts* C44, responding to ApG as primer but not to globin mRNA (Mahy *et al.*, 1981 a, b).

These assays were designed to detect differences in *in vitro* virion RNA transcriptase activities of mutants at 40.5 °C compared to 31 °C, but some of the *ts* mutants analyzed displayed 2—3 fold less transcriptase activity at 31 °C than the wild-type. Previously Mowshowitz and Ueda (1976), working with the New York-Tokyo mutants, had shown that two group I mutants, with lesions in RNA segment 1 (*ts* NYT1 and *ts* NYT6), were temperature-sensitive for virion transcriptase activity *in vitro;* however these studies were all carried out below the permissive temperature for mutant growth *in vivo.* By use of 1 mM guanosine (McGeoch and Kitron, 1975) as primer, virion RNA transcriptase activities were measured at 25 °C, 28 °C, and 31 °C; over this range, the activity of ts^+ virus was stable, but the activities of *ts* NYT1 and *ts* NYT6 decreased to about 20 per cent at 31 °C compared to 25 °C (Mowshowitz and Ueda, 1976). None of the mutants studied from other recombination groups displayed such a defect. In addition, Mowshowitz and Ueda (1976) found that the temperature-sensitivity of the virion transcriptase was reversible, and not due to increased thermolability, as measured by infectivity of the group I mutants following pre-incubation at the non-permissive temperature. Confirmation of these results has recently been provided by Ulmanen *et al.* (1983), who used lysolecithin-disrupted virus cores (Rochovansky, 1976) rather than detergent-disrupted virus to analyze the virion transcriptase activity of *ts* NYT1 and *ts* NYT6. Using cores, similar transcription rates were found for ts^+ virus at 39.5 °C and 33 °C, whereas with both mutants *in vitro* transcription was only about 15% at 39.5 °C compared to 33 °C. The assay mixture in these experiments included alfalfa mosaic virus RNA 4 as a capped mRNA primer molecule, and the *ts* reaction step appeared to be recognition of the 5′-terminal cap structure on the primer, in agreement with the known function of PB_2 in cap-binding (Ulmanen *et al.*, 1983).

It is not clear why some *ts* mutants display reduced virion RNA transcriptase activity at the permissive temperature, since they must function at this temperature in the infected cell. However, the results of Ulmanen *et al.* (1983) suggest that the lipid envelope may be inhibitory, and perhaps once the virion is uncoated, the *ts* polymerase functions normally in the infected cell at the permissive temperature. Alternatively, it may be that a virion RNA transcriptase activity which is 5% of that in wild-type virus is still sufficient to initiate a normal infection. Bean and Simpson (1975), working with the New York WSN strain mutants, found one (*ts* NY24) in which virion RNA transcriptase activity was barely detectable at 37 °C and never more than 5% of that of wild-type virus when assayed at 32 °C, the permissive temperature. Unfortunately, it is not known which RNA segment carries the mutation lesion in *ts* NY24. The transcriptase activity of *ts* NY24 could not be increased by variations in the reaction conditions, including alterations in divalent cations, detergent concentration, or temperature. The RNA-dependent RNA polymerase activity induced in the microsomal fraction of cells infected with this mutant was also greatly reduced compared to ts^+ virus-infected cells (Bean and Simpson, 1975). Reduced levels of virion RNP have also been found in cells infected with this mutant, despite the fact that normal yields of infectious virus were observed (Gregoriades and Hirst, 1976).

The virion RNA transcriptase acitivities of the fowl plague (Weybridge) mutants isolated in Moscow were analyzed *in vitro* at 38 °C and 42 °C by Ghendon *et al.* (1973). The only RNA segment 1 mutant tested was *ts* M29; the reduction in enzyme activity at 42 °C compared to 36 °C was similar to that observed with wild-type virus (7—20%). It was concluded that *ts* M29 does not possess a defective virion transcriptase; a repeat of this experiment reported in a later paper gave further support to this conclusion (Ghendon *et al.*, 1975).

b) RNA-Dependent RNA Polymerase Activity of Infected Cell Extracts

RNA-dependent RNA polymerase activity can be detected in both the nuclear and microsomal fractions of influenza virus-infected cells (Scholtissek and Rott, 1969 a; Skehel and Burke, 1969; Mahy and Bromley, 1970; Hastie and Mahy, 1973; Mahy *et al.*, 1975; Simpson and Bean, 1975). This enzyme activity has been used by the Giessen group as a marker for phenotypic alterations in virus-induced RNA synthesis of fowl plague virus Rostock mutants (Scholtissek *et al.*, 1974; Scholtissek and Bowles, 1975; Scholtissek, 1979). The RNA-dependent RNA polymerase activity is measured *in vitro* using reaction conditions similar to those used for the virion transcriptase, except that detergents are not required. The RNA synthesized by the microsomal fraction *in vitro* is complementary to virion RNA (Scholtissek, 1969; Hastie and Mahy, 1973); as this enzyme activity increases relatively late in the multiplication cycle, it probably derives from RNA transcriptase complexes destined for incorporation into mature virions (Mahy *et al.*, 1975; Simpson and Bean, 1975). It is not clear what the RNA-dependent RNA polymerase activity of the nuclear fraction represents. This enzyme activity can be detected early in the multiplication cycle, it synthesizes mainly complementary RNA *in vitro*, and it has requirements similar to the microsomal enzyme for optimum *in vitro* activity (Hastie and Mahy, 1973). Neither the nuclear nor the microsomal RNA polymerase can be detected earlier than 2 hours post-infection, and most analyses have been carried out at 5 hours post-infection. Mutants that are *ts* for RNA polymerase induction may therefore have defects at any one of a large number of steps in the multiplication cycle, not all of which are concerned with RNA synthesis. Not surprisingly, Scholtissek and Bowles (1975) found polymerase negative mutants in all 6 of the recombination groups which they studied.

Ten group I mutants (with *ts* lesions in RNA segment 1) were analyzed for induction of RNA-dependent RNA polymerase activity in nuclear and microsomal fractions 5 hours after infection at 33 °C or 40 °C (Scholtissek and Bowles, 1975). Eight of the mutants were polymerase negative (*ts* G3, 8, 115, 117, 132, 206, 254, and 293) and two were polymerase positive (*ts* G121 and 290). Negative mutants induced less than 5% activity at 40 °C compared to 33 °C. A more detailed study of induced RNA polymerase activity was made with *ts* G3. After infection at 33 °C, the polymerase activity induced by this mutant was not destroyed *in vivo* if cells were shifted up to 40 °C at 4 hours post-infection, then analyzed at 5 or 6 hours post-infection (Scholtissek *et al.*, 1974). In addition, the RNA polymerase activity induced by *ts* G3 at 33 °C was as stable as the wild-type virus-induced enzyme when

tested *in vitro* at 40 °C (Scholtissek *et al.*, 1974), although the wild-type enzyme is inactivated quite rapidly *in vitro* above 38 °C (Scholtissek and Rott, 1969 b).

c) Virus-Specific RNA Synthesis *in vivo*

In influenza virus-infected cells, three classes of virus-specific RNA can be detected (Skehel and Hay, 1978 a, b; Hay *et al.*, 1980, 1982; Krug this volume). The first of these classes, A(+) cRNA, comprises the functional messenger RNAs (Glass *et al.*, 1975) and is synthesized in infected cells during "primary transcription" (Bean and Simpson, 1973) in the absence of virus-specific protein synthesis. A(+) cRNAs are incomplete genome transcripts, lacking about 16 nucleotides complementary to the 5′ end of virion RNA (Skehel and Hay, 1978 a, b; Robertson *et al.*, 1981) and with additional 3′-terminal polyadenylate sequences and heterogeneous capped extensions of 10 to 15 nucleotides at their 5′ termini. The latter are apparently derived from the termini of cellular mRNAs or their precursors during priming of transcription of viral A(+) cRNAs (Krug *et al.*, 1979; Caton and Robertson, 1980; Dhar *et al.*, 1980; Krug, 1981). The second class of virus-specific RNA, A(−) cRNAs, comprises complete genome transcripts with 5′-terminal pppA, which do not require "priming" for their synthesis, and are considered to be the templates for new genomic RNA synthesis during replication (Hay *et al.*, 1982). The third class of virus-specific RNA synthesized in infected cells comprises negative strand virion genomic RNAs (vRNA). Neither A(−) cRNAs nor vRNAs are synthesized in infected cells under conditions of primary transcription, when protein synthesis is blocked by addition of cycloheximide. Early studies of RNA synthesis in cells infected with *ts* mutants measured cRNA and vRNA but did not distinguish the two classes of cRNA. More recently, all three classes have been analyzed using methods which measure either total accumulation of virus-specific RNA by hybridization to radiolabelled probes (Barrett *et al.*, 1979; Mark *et al.*, 1979; Taylor *et al.*, 1977) or newly synthesized virus-specific RNA pulse-labelled in infected cells and hybridized to an excess of unlabelled vRNA or cRNA (Hay *et al.*, 1977 a; Mowshowitz, 1981) or to single-stranded influenza virus-specific DNA cloned in phage M13 (Thierry and Danos, 1982).

Early studies. Early studies of the RNA phenotype of influenza virus mutants analyzed ^{3}H-uridine incorporation in infected cells to which actinomycin D had been added at 3–4 hours post-infection (Ghendon *et al.*, 1973; Sugiura *et al.*, 1975). Although influenza virus replication is inhibited by actinomycin D (reviewed by Barry and Mahy, 1979), addition of the drug relatively late in the growth cycle at concentrations which suppress host cell RNA synthesis does not completely inhibit virus-specific RNA syntheses. In particular, whilst the synthesis of A(+) cRNA, which occurs early in infection, is actinomycin-sensitive, the synthesis of A(−) cRNA and vRNA, which do not require "priming" by host cell mRNAs (Hay *et al.*, 1982), are relatively resistant to the drug (Scholtissek and Rott, 1970; Pons, 1973; Staroff and Bukrinskaya, 1981).

Using this approach, Ghendon *et al.* (1973) found that *ts* M29, with a *ts* lesion in RNA segment 1, was negative for actinomycin D-resistant RNA synthesis at the non-permissive temperature. They also carried out temperature shift experiments

using *ts* M29 (Ghendon *et al.*, 1975). Infected cells were incubated at the permissive temperature (36 °C) for 3 hours, then actinomycin D was added and at 3.75 hours post-infection the cells were labelled with ^{3}H-uridine for a further 2 hours at either 36 °C or 42 °C. No RNA synthesis could be detected in *ts* M29 infected cells when ^{3}H-uridine labelling was performed at the non-permissive temperature (Ghendon *et al.*, 1975). In a more detailed study, Sugiura *et al.* (1975) used a similar method to compare RNA synthesis by *ts* mutants and wild-type WSN virus at the non-permissive temperature. After incubating infected cultures at 39.5 °C for 3 hours, actinomycin D was added, followed one hour later by ^{3}H-uridine, in the continuous presence of actinomycin D, for a further 2 hours (*i.e.* 4 to 6 hours post-infection). After subtracting background incorporation of ^{3}H-uridine into mock-infected cultures, the incorporation of ^{3}H-uridine by *ts* mutant-infected cultures was expressed as a percentage of the incorporation observed during infections with wild-type virus. Using this assay, seven RNA segment 1 mutants (*ts* NYT 1, 6, 8, 54, 58, 63, and 65) were completely negative for RNA synthesis, and one (*ts* NYT 55) gave 16.7% incorporation. *Ts* NYT 55 was not leaky under the experimental conditions used, and so must be considered RNA positive by this analysis, and an exception to the other segment 1 mutants. Sugiura *et al.* (1975) attempted to determine whether cRNA or vRNA synthesis was primarily affected by keeping two sets of cultures at the permissive temperature for the first 4 hours post-infection, before adding actinomycin D. Then while one set remained at the permissive temperature the other was shifted up to the non-permissive temperature, and after 1 hour all cultures were labelled for 2 hours with ^{3}H-uridine and the levels of RNA synthesis at the two temperatures were compared. Six segment 1 mutants tested were negative for RNA synthesis after shift-up (*ts* NYT 54 and 55 were not included in this series).

Analysis of RNA phenotypes by hybridization. Ghendon *et al.* (1975) measured virus-specific cRNA synthesis in cells infected with *ts* M29 by labelling with ^{3}H-uridine from 1—2.5 or from 2.5—4 hours post-infection and determining RNase-resistant radioactivity of RNA extracted from infected cells after annealing to an excess of virion RNA. Approximately 3% of ^{3}H-uridine-labelled RNA was RNase-resistant in cells infected at 36 °C compared to 1% in cells infected at 42 °C. Clearly there are difficulties in interpreting these results, but the authors concluded that synthesis of cRNA is reduced in *ts* M29-infected cells at 42 °C. This mutant has recently been analyzed in more detail by Ghendon *et al.* (1982 a) using the procedures of Hay *et al.* (1977 a). RNA extracted from infected cells, ^{3}H-uridine-labelled from 2.25—3 hours post-infection, was annealed with excess unlabelled virion RNA. The A(+) and A(−) hybrid molecules were separated by LiCl fractionation and the former incubated with nuclease S_1, then the double-stranded RNAs were analyzed by electrophoresis in polyacrylamide gels. These experiments showed that the level of mRNA synthesis in cells infected with *ts* M29 at 42 °C was less than 5% of that in cells infected at 36 °C, whereas mRNA synthesis by wild-type virus was similar at either temperature. However, synthesis of A(+) cRNA in the presence of cycloheximide (primary transcription) was similar at either temperature in *ts* M29-infected cells. Temperature shift experiments were performed in which, after 2 hours incubation at 36 °C, cycloheximide was added and the cells shifted up to 42 °C. Under these conditions, A(+) cRNA synthesis in *ts* M29 infected cells ceased,

and it was concluded that *ts* M29 has a defect in "secondary" mRNA synthesis. A corollary of this finding is that "primary" and "secondary" syntheses of mRNA involve different enzymes or enzyme complexes. In addition to the defect in "secondary" mRNA synthesis, *ts* M29 was found to be defective in the syntheses of both A(−) cRNA and vRNA (Ghendon *et al.*, 1982 a).

The WSN mutants of Sugiura *et al.* (1975) were further analyzed for virus-specific RNA synthesis by Krug *et al.* (1975) and Mowshowitz (1981) during infection of MDBK cells. The synthesis of cRNA was analyzed in the earlier experiments by hybridizing RNA extracted from infected cells to ^{3}H-adenosine-labelled virion RNA after boiling in low salt to dissociate virus-specific double-stranded forms. As the specific activity of the virion RNA was only about 10^4 counts/min/μg, the sensitivity of this procedure was relatively low, and primary transcription could not be detected. However, one group I mutant (*ts* NYT 6) induced the synthesis of no detectable cRNA at 39.5 °C compared to 33 °C, and if cells were left at the permissive temperature for 5 hours after *ts* NYT 6 infection and then shifted up to 39.5 °C, no further increase in the amount of cRNA could be detected (Krug *et al.*, 1975). Later experiments established that the *ts* mutation in *ts* NYT 6 was in RNA segment 1 (Palese *et al.*, 1977 b).

Mowshowitz (1981) measured newly synthesized RNA in WSN mutant-infected MDBK cells by pulse-labelling at various times with ^{3}H-uridine. Total cRNA was then determined by hybridization to unlabelled virion RNA and total vRNA was determined by hybridization to unlabelled mRNA isolated from infected cells. To determine "template" cRNA, "total" ^{3}H-uridine-labelled cRNA hybridized to unlabelled vRNA was passed twice through a column of benzoylated DEAE cellulose, which binds nucleic acids with single-stranded regions. The material which passed through such a column was analyzed for RNase-resistant radioactivity as a measure of template cRNA, which represented roughly a third of the total cRNA synthesized during wild-type virus infection. *Ts* NYT 1 (group I) was found to be grossly deficient both in cRNA and vRNA synthesis when infected cells, after 4 hours incubation at 33 °C, were shifted to 39.5 °C and pulse-labelled with ^{3}H-uridine for 1 hour.

RNA synthesis by the Giessen mutants of fowl plague virus was similarly analyzed by pulse-labelling infected cells with ^{3}H-uridine from 4.5 to 6 hours post-infection and hybridizing total labelled RNA to unlabelled vRNA or cRNA (Scholtissek *et al.*, 1974). Two mutants with lesions in RNA segment 1 (*ts* G3 and *ts* G132) were each found to be negative for RNA synthesis in cells kept at the non-permissive temperature throughout infection, or if cultures were shifted from the permissive to the non-permissive temperature at 3.5 hours before pulse-labelling (Scholtissek *et al.*, 1974; Scholtissek and Bowles, 1975).

Preliminary studies with the Cambridge fowl plague virus mutants showed that one RNA segment 1 mutant, *ts* C17, was defective in synthesis of all classes of virus-specific RNA at the restrictive temperature (Barry and Mahy, 1979). Analysis of the synthesis of the three classes of RNA was carried out as described by Barrett *et al.* (1979), by hybridizing specific radiolabelled probes to unlabelled RNA extracted from the mutant-infected cells. vRNA was measured by hybridization to a ^{3}H-cDNA probe prepared by reverse transcription of virion RNA in the presence of

an oligo(dG) primer. This probe contained sequences from the entire genome (Barrett *et al.*, 1979). Complementary RNA was detected by hybridization to I^{125}-labelled virion RNA; A(+) cRNA and A(−) cRNA fractions were preselected by passing the total RNA through an oligo(dT)-cellulose column. Primary transcription was measured by incubating infected chick embryo fibroblast cells in the presence of cycloheximide throughout infection and analyzing RNA extracted from the cells at 5 hours post-infection at either 34 °C or 40.5 °C (Mahy *et al.*, 1981 a, b; Penn, 1983). All nine RNA segment 1 mutants studied (*ts* C3, 5, 9, 17, 21, 25, 33, 44, and mN5) synthesized reduced amounts of cRNA at 40.5 °C compared to 34 °C, whereas in wild-type virus-infected cells the amounts of cRNA were similar at either temperature. The concentration of virus-specific A(+) cRNA (ng per μg cell RNA) in RNA segment 1 mutant infected cells at 40.5 °C varied from 29% to 53% of that found in similar infections at 34 °C. Two of these mutants (*ts* C17 and mN5) were examined for their ability to induce A(−) cRNA and vRNA synthesis during infections at the two temperatures in the absence of cycloheximide. Synthesis of both types of RNA was drastically reduced at 40.5 °C compared to 34 °C (Mahy *et al.*, 1981 a). Temperature shift-up experiments were carried out with mutant *ts* C mN5. Infection was allowed to proceed 2 hours at 34 °C before shifting to 40.5 °C, and continuing incubation for a further 2 hours before RNA extraction. Under these conditions, the synthesis of both A(−) cRNA and vRNA reached similar levels to those found in cells kept at 34 °C throughout the 4 hour infection period. It was concluded that *ts* C mN5 is defective in mRNA synthesis but not in template or vRNA synthesis (Mahy *et al.*, 1981 b).

d) Virus-Specific Polypeptide Synthesis

Despite the defects in virus-specific RNA synthesis which are found during infections at the non-permissive temperature with RNA segment 1 mutants, few alterations in virus-specific polypeptide synthesis have been reported. Ghendon *et al.* (1973) found an almost normal pattern of polypeptide synthesis by pulse-labelling cells infected at 42 °C for 5 hours with *ts* M29. Similarly, Sugiura *et al.* (1975) found almost normal amounts of RNP, HA and NA at the restrictive temperature after 10 hours infection with *ts* NYT 8, 54, 55, 58, 63, and 65, although infection with two other segment 1 mutants, *ts* NYT 1 and 6, resulted in greatly reduced synthesis of these products. Sugiura *et al.* (1975) speculated that the RNA-negative mutants which synthesized virus-specific proteins might do so because a small amount of mRNA is transcribed due to a low level of leak, or during unamplified primary transcription, which might be sufficient to code for detectable amounts of protein.

Two of the Giessen segment 1 mutants of fowl plague virus Rostock were examined for their ability to induce virus-specific protein synthesis. If cells were maintained at 40.5 °C throughout infection, no virus-specific protein synthesis could be detected with *ts* G3 (Scholtissek, personal communication; Penn, 1983). However, if infected chick embryo fibroblast cell cultures were incubated at the permissive temperature (33°) for 3 hours (*ts* G3) or 4 hours (*ts* G293) before shift-up to 40 °C, followed by labelling for 1 hour with variious ^{3}H-amino-acids, both mutants induced approximately normal amounts of all virus-specific polypeptides,

though it was noted that *ts* G293-infected cells incorporated three to five times the normal amount of radioactivity into the P protein region of the gel (Scholtissek *et al.*, 1974; Scholtissek and Bowles, 1975). The significance of this observation is unclear, although over-production of P protein mRNA during incubation at the permissive temperature seems likely.

Recently the fowl plague virus Rostock *ts* mutants isolated in Cambridge have been analyzed for their ability to induce virus-specific protein synthesis in chick embryo fibroblast cells by high resolution polyacrylamide gel electrophoresis of ^{35}S-methionine-labelled polypeptides followed by autoradiography (Inglis *et al.*, 1976). Using this approach, three distinct stages of virus-specific protein synthesis are characteristically seen during normal productive infection: an immediate-early (primary transcription) stage when all polypeptides are synthesized at equivalent rate, an early stage (up to 2 hours post-infection at 37 °C) when amplification of NP and NS_1 protein synthesis occurs, and a late stage when amplification of HA, M, and NS_2 polypeptides is readily observed (Inglis *et al.*, 1978; Inglis and Mahy, 1979). The early and late stages were similar in cells infected at 34 °C (the permissive temperature) but at 40.5 °C the infection progressed more rapidly, and a late protein synthesis pattern was established by 2 hours post-infection (Mahy *et al.*, 1981 a; Penn, 1983). Nine RNA segment 1 mutants were studied. In each case, the pattern of virus-induced protein synthesis at 34 °C was similar to that of wild-type virus, and a late protein synthesis pattern developed by 4—6 hours post-infection. At 40.5 °C, two mutants (*ts* C21 and *ts* C33) induced a protein synthesis pattern similar to ts^+ virus. With six of the mutants, the characteristic late polypeptide pattern did not develop at 40.5 °C; M was synthesized at the same rate as NS_1, and synthesis of HA was not detected by ^{35}S-methionine labelling, although synthesis of the spliced gene product, NS_2, did occur. The exception was *ts* C3, with which amplification of NS_1 and NP protein synthesis did not precede that of M synthesis, and the onset of protein synthesis was late after infection. Synthesis of NS_2 polypeptide was not observed in *ts* C3-infected cells at 40.5 °C. An attempt was made to quantitate the level of protein synthesis in these *ts* mutant-infected cells by Kenacid blue staining (Penn, 1983). This confirmed that less virus-specific protein accumulated at 40.5 °C with all segment 1 mutants studied. These results with RNA segment 1 mutants are not sufficiently clear-cut to lead to firm conclusions on the role of protein PB2 during infection except that there may be a specific requirement for PB2 in controlling the switch from early to late protein synthesis.

e) Summary

A large number of *ts* mutants with lesions in RNA segment 1 (coding for PB2) have been examined for phenotypic alterations at the restrictive temperature (Table 5). Both the Cambridge and the Giessen collections of fowl plague virus (Rostock) mutants contain more mutants in the group with lesions in RNA segment 1 than in any other recombination group. As would be predicted from the role of PB2 in cap-binding, all show defects in RNA transcription, and eleven of the mutants have been found defective in recognition of capped primer mRNA molecules. The variable results found in protein synthesis studies indicate, however, that mRNA

Table 5. *Phenotype at the restrictive temperature of influenza A mutants having a* ts *lesion in RNA segment 1*

Mutant	RNA transcriptase		Virus-specific RNA			Virus-specific polypeptides		Comments
	Virion	Induced	Total[a]	A(+)[b] cRNA	A(−)cRNA and vRNA[b]	Early	Late	
C3	−			±		±	±	Onset of all polypeptide synthesis delayed, and NS_2 not synthesized
C5	−			±		+	−	
C9	−			±		+	−	
C17	−			±	−	+	−	Virion transcriptase not stimulated
C21	−			±		+	+	by capped (globin) mRNA with any of
C25	−			±		+	−	the Cambridge mutants
C33	−			±		+	+	
C44	−			±		+	−	
CmN5	−			±	−	+	−	A(−) cRNA and vRNA normal after shift-up at 2 hrs.
G3		−	−		−	−	−	Late protein synthesis normal after shift-up at 3 hrs.
G8		−						
G115		−						
G117		−						
G121		+						
G132		−	−		−			
G206		−						
G254		−						
G290		+						
G293		−						Late protein synthesis measured after shift-up at 4 hrs.
M29	+		−	±	−		+	Secondary mRNA synthesis is defective, but not primary
NYT1	−		−	−	−		−	Defective in binding of capped RNA primer molecule
NYT6	−		−	−			−	Defective in binding of capped RNA primer molecule
NYT8			−				+	
NYT54			−				+	Protein synthesis of all NYT mutants was
NYT55			+				+	assessed as accumulation of virus-specific
NYT58			−				+	RNP, HA and NA
NYT63			−				+	
NYT65			−				+	

[a] Measured by ^{3}H-uridine labelling in presence of actinomycin D.
[b] Measured by hybridization to specific probes.

synthesis is not completely blocked *in vivo,* and only one mutant which has been rigorously examined, *ts* G3, is completely negative in this respect. Protein synthesis by many of the Cambridge mutants differs from that of wild-type virus in that synthesis of HA is not detected at the restrictive temperature, and amplification of M relative to NS1 synthesis does not occur. These mutants appear to be defective in the regulation of transcription; in this respect it is interesting that, at the restrictive temperature, *ts* M29 carries out primary, but not secondary, mRNA synthesis. This suggests that, in addition to its role in recognition of capped RNA molecules for use as primers, PB2 is involved in control of transcription at later stages in the replication cycle.

Two of the Cambridge RNA segment 1 mutants (*ts* C21 and *ts* C33) display reduced mRNA synthesis, but no detectable differences in virus-specific polypeptide synthesis at the restrictive temperature. It is not clear why such mutants do not form plaques at the restrictive temperature, although there could be an assembly defect not so far detected. However, some mutations might not totally abolish virus growth, but so reduce it that plaques are not formed within the time of a standard plaque assay. Consistent with this notion, Almond (1978) found that *ts* C21 and *ts* C33 were both leaky, and capable of producing some infectious (but still *ts*) virus after a single replication cycle at 40.5 °C.

3. Mutants with ts *Lesions in RNA Segment 2*

RNA segment 2 encodes a basic polypeptide (PB1) which in the PR8 virus strain has a molecular weight of 86,500 (Winter and Fields, 1982) and is the slowest migrating virus-specific polypeptide in polyacrylamide gel analysis of all influenza virus strains so far examined. Recently, two pieces of biochemical evidence have established that a function of PB1 is to initiate virion RNA transcription by catalyzing the elongation of the primer molecule. Influenza virion RNA segments all contain at their 3′-termini the sequence UpCpGpUpU (Skehel and Hay, 1978 a; Robertson, 1979; Desselberger *et al.,* 1980). The initiation of transcription, with capped RNAs as primers, occurs through the incorporation of a guanosine onto the 3′ termini of the primer fragments generated by the virion endonuclease activity (Plotch *et al.,* 1981). When influenza (WSN strain) virion cores were incubated for 30 min at 31 °C with (α-^{32}P)-GTP in the presence of globin mRNA as primer, then irradiated for 60 min at 0 °C with 3×10^5 ergs/mm^2 UV light, the PB1 protein became specifically crosslinked to the labelled guanosine residue (Ulmanen *et al.,* 1981). If, instead of a capped mRNA fragment, the primer is the dinucleotide ApG (McGeoch and Kitron, 1975), detergent-disrupted influenza (fowl plague Rostock) virions catalyze the reaction ApG + pppC → ApGpC + ppi in the presence of (α-^{32}P)-CTP as the sole available nucleotide, and during this reaction PB1 forms a covalent complex with the labelled cytidylate residue (Horisberger, 1982). It thus appears that PB1 carries a catalytic site for the addition of guanosine or cytosine to a primer molecule during initiation of transcription, though it is not yet established whether this protein also catalyzes chain elongation. Like PB2, PB1 specifically migrates into the cell nucleus during the early stages of virus replication (Briedis *et al.,* 1981).

a) Virion Transcriptase Activity

Early studies of the *in vitro* RNA transcriptase activity of RNA segment 2 (group III) mutants from the Sugiura collection gave equivocal results (Mowshowitz and Ueda, 1976). Both *ts* NYT15 and *ts* NYT101 gave variable results during *in vitro* transcriptase assays measured over a range of temperatures between 25 °C and 36 °C (*i.e.* below the restrictive *ts* mutant temperature). Mowshowitz (1978) made a more definitive study of *ts* 101 by employing dinucleotide primer molecules (McGeoch and Kitron, 1975) in the transcriptase reaction. The concentration of the dinucleotide ApG was varied in reactions catalyzed by detergent-disrupted ts^+ or *ts* 101 at 35 °C or 28 °C. As the ApG concentration was decreased, *ts* NYT 101 became increasingly temperature-sensitive for RNA transcription, and it appeared that ts^+ and *ts* NYT101 possessed different affinities for the ApG primer at the higher temperature. It was concluded that the RNA segment 2 product is required for the initiation of cRNA synthesis (Mowshowitz, 1978). A later study by Horisberger (1982) confirmed that *ts* NYT101 was temperature-sensitive even at 28 °C if low concentrations of ApG were employed in the reaction mixture. The synthesis of ApGpC occurred at similar rates in reactions catalyzed by *ts* NYT101 or ts^+ virus if the ApG concentration was kept at 0.4 mM, but at 0.05 mM ApG, the mutant enzyme activity was rapidly inactivated (Horisberger, 1982).

Comparison of the transcriptase activity of fowl plague virus mutants isolated in Moscow showed that with one RNA segment 2 mutant, *ts* M131, *in vitro* transcriptase activity was reduced by about 75% at 42 °C compared to 36 °C, whereas ts^+ virus activity was reduced by only 15% at the higher temperature (Ghendon *et al.*, 1973, 1975). A more recent study by Nichol (1980) showed that *ts* M131 was also defective in ApG-primed RNA transcription *in vitro*. Direct recombination analysis between the *ts* M mutants of fowl plague virus Weybridge and the *ts* C mutants of fowl plague virus Rostock confirmed that *ts* M131 is defective in the RNA segment equivalent to fowl plague virus Rostock RNA segment 2 (Ghendon *et al.*, 1982 d).

Nichol *et al.* (1981) found that the virion RNA transcriptase activity of two fowl plague virus Rostock mutants, *ts* C1 and *ts* C15 was only slightly reduced at 40.5 °C compared to ts^+ virus, and responded similarly to ApG or globin mRNA added as a primer. Thus, the PB_1 polypeptide does not appear to play a role in recognition of the capped primer molecule.

b) RNA-Dependent RNA Polymerase Activity of Infected Cell Extracts

Two Giessen mutants of fowl plague virus Rostock belonging to recombination group III, *ts* G90 and *ts* G93, were examined for their ability to induce RNA-dependent RNA polymerase activity in the cytoplasm or nucleus of cells incubated at 33 °C or 40 °C for 5 hours post-infection (Scholtissek and Bowles, 1975). Polymerase activity was not detected in the cytoplasm of *ts* mutant-infected cells at either temperature, but at 33 °C polymerase activity accumulated in the nuclear fraction. At 40 °C, the nuclei of *ts* mutant-infected cells possessed about 20% of the RNA-dependent RNA polymerase activity found in ts^+ infected cells. The accumulation of RNA-dependent RNA polymerase in cell nuclei at 33 °C following infection with *ts* G90 or *ts* G93 was interpreted as a defect in the transport of the

polymerase complex from nucleus to cytoplasm, and the product of RNA segment 2 was described as the "transport protein" (Scholtissek, 1978). It is interesting in this connection that one *ts* mutant studied by Mackenzie and Dimmock (1973), *ts* 20, was defective in the transport of RNP from the nucleus during infection at the restrictive temperature. Despite this "transport" defect, virus-specific RNA synthesis by the Giessen mutants was not impaired at the restrictive temperature (Scholtissek, 1979). Heller and Scholtissek (1980) later described two further mutants of the group (*ts* G18 and *ts* G236), which induced no RNA-dependent RNA polymerase activity in cell extracts at the restrictive temperature (Scholtissek and Bowles, 1975).

c) Virus-Specific RNA Synthesis *in vivo*

Early studies of fowl plague virus Weybridge mutants showed that *ts* M131 was defective in cRNA synthesis at 42 °C (Ghendon *et al.*, 1973, 1975). Temperature shift experiments were carried out in which *ts* M131-infected cells were kept at 36 °C for 3 hours, then treated with actinomycin D and shifted up to 42 °C before labelling with ^{3}H-uridine. Under these conditions (which should detect mainly vRNA synthesis) no labelling of virus-specific RNA or detectable formation of virus-specific RNP structures were found (Ghendon *et al.*, 1975). In a more detailed investigation of *ts* M131, Ghendon *et al.* (1982 a) found impairment in the syntheses of both A(−) cRNA and vRNA at the restrictive temperature. A(+) cRNA synthesis was not affected if measured in the presence of cycloheximide ("primary transcription"), but was greatly reduced at 42 °C in the absence of the protein synthesis inhibitor. The latter reduction was considered to be secondary to the defect in vRNA synthesis at the restrictive temperature, and not to result from a defect in A(+) cRNA synthesis by *ts* M131 (Ghendon *et al.*, 1982 a). Such a dependence of mRNA synthesis on newly synthesized vRNA would be consistent with the patterns of mRNA and vRNA synthesis which have been observed during fowl plague virus replication by Smith and Hay (1982). On the other hand, the reductions in both A(−) cRNA and vRNA synthesis in *ts* M131-infected cells at the restrictive temperature appeared to be direct consequences of the temperature-sensitivity of the RNA replicase (Ghendon *et al.*, 1982 a), inferring that PB1 is a component of this enzyme.

The Giessen fowl plague virus mutants have not been studied in as much detail, but the two mutants with a "transport" defect, *ts* G90 and *ts* G93, produced almost normal amounts of vRNA and cRNA at the restrictive temperature (Scholtissek and Bowles, 1975), as did *ts* G18 (Heller and Scholtissek, 1980; Scholtissek, 1979). Another of the Giessen mutants with a defect in RNA segment 2 (Heller and Scholtissek, 1980), was, however, reported to be negative for virus-specific RNA synthesis at the restrictive temperature (Scholtissek and Bowles, 1975).

Virus-specific RNA synthesis was studied with one of the Cambridge segment 2 mutants, *ts* C15 (Mahy *et al.*, 1981 a). Primary transcription in the presence of cycloheximide was reduced by about 50% at 40.5 °C compared to 34 °C; this reduction was similar to that seen with RNA segment 1 mutants. Secondary transcription

and vRNA synthesis were more severely affected, however. Analysis of A(−) cRNA and vRNA was carried out after incubating *ts* C15-infected cells for 2 hours at 34 °C, then shifting to 40.5 °C for a further 2 hours. Under these conditions, similar amounts of A(−) cRNA were detected in temperature-shifted cultures compared to cultures held at 34 °C throughout the 4 hour period. However, very little vRNA synthesis was detected in shifted cultures, suggesting that PB1 is directly involved in vRNA synthesis (Mahy *et al.*, 1981 a). It should be noted that these experiments measured only the total accumulated amount of each RNA species; as most A(−) cRNA appears to have been synthesized by 2 hours at 34 °C in this system, the failure to accumulate vRNA between 2 and 4 hours at 40.5 °C is consistent with the suggestion by Ghendon *et al.* (1982 a) that PB1 is a component of the replicase.

Sugiura *et al.* (1975) reported that three group III (RNA segment 2) mutants, *ts* NYT5, 15, and 101, were all RNA negative when assayed for ^{3}H-uridine incorporation in the presence of actinomycin D added 3 hours post-infection. In a more detailed hybridization analysis, Krug *et al.* (1975) found that *ts* NYT5 synthesized some cRNA at the restrictive temperature, but *ts* NYT15 and 101 were completely negative. In temperature shift experiments, *ts* NYT15-infected cells were maintained at 33 °C for 5 hours then shifted to 39.5 °C. Under these conditions cRNA accumulation ceased after shift-up, and it was concluded that protein P1 (PB1) was essential for cRNA synthesis (Krug *et al.*, 1975; Palese *et al.*, 1977 b). This was confirmed in a later kinetic analysis which measured newly synthesized cRNA (Mowshowitz, 1981).

One of the Bethesda mutants of the Hong Kong strain of influenza virus, *ts* B315, has been analyzed by annealing to single-stranded DNA probes of a haemagglutinin gene fragment prepared in phage M13 (Thierry and Danos, 1982). DNA probes were prepared in either orientation to permit detection of cRNA or vRNA sequences, and prior selection of poly(A)-containing RNA by oligo(dT) cellulose chromatography allowed separate determination of A(+) and A(−) cRNA fractions. Cells were labelled with ^{3}H-adenosine at various times post-infection, and the RNA hybridized to filters containing the respective unlabelled DNA probes. In *ts* B315-infected cells, the rates of accumulation of all three RNA species studied were reduced at 39 °C to 20—50% of the rates seen at 32 °C. Temperature-shift experiments showed that the primary defect in *ts* B315 was a failure to synthesize A(−) cRNA, resulting in defects both in vRNA synthesis and in amplification of A(+) cRNA (Thierry and Danos, 1982). These results are thus consistent with a role for PB_1 in replicase activity.

d) Virus-Specific Polypeptide Synthesis

Few studies of virus-induced protein synthesis in cells infected with RNA segment 2 mutants have been reported. Ghendon *et al.* (1973) found a severe reduction in all virus-induced protein synthesis in *ts* M131-infected cells held at 42 °C for 5 hours before labelling with radioactive amino-acids followed by polyacrylamide gel electrophoresis. In a later study, they also reported that *ts* M131-infected cells held at 42 °C displayed no characteristic ultrastructural changes such as were seen at 36 °C,

and no virus maturation was observed at the cell membranes (Anisimova *et al.*, 1980).

Sugiura *et al.* (1975) monitored the production of RNP, HA and NA over a 10 hour incubation at 39.5 °C, the restrictive temperature. All group III *ts* mutants studied (*ts* NYT5, 15 and 101) were grossly deficient in production of these virus-specific components. One of the Cambridge mutants, *ts* C15, studied by ^{35}S-methionine-labelling at 2 and 4 hours post-infection, was found to induce the synthesis of all virus-specific polypeptides during incubation at 40.5 °C (Mahy *et al.*, 1981 a). More detailed analysis of polypeptide synthesis by this and other group II mutants has not so far been carried out.

e) Summary

Few RNA segment 2 mutants have been studied in any detail (Table 6), but recent results with *ts* NYT101 have confirmed that PB1 is essential for the initiation of RNA transcription. There is also evidence from five mutants that the synthesis of vRNA is also PB1-dependent, though at what stage is not yet clear. One possibility is that initiation of vRNA synthesis also requires PB1; another is that transport of the polymerase complex from nucleus to cytoplasm is essential for vRNA synthesis, and this requires PB1. The two Giessen mutants, G90 and G93, with which the RNA polymerase complex is held up in the nucleus at the restrictive temperature, led Scholtissek (1978) to refer to this segment at the "transport gene", and its product as P_{tra}. Whilst this may be one function of the RNA segment 2 product, it is clearly not the only function, since it is unlikely to be related, for example, to the initiation process which has been shown to be *ts in vitro*. It would thus seem safer, until more information as to function becomes available, to refer to the segment as 2 and the product as PB1.

4. Mutants with ts *Lesions in RNA Segment 3*

RNA segment 3 encodes an acidic polypeptide (PA) which in the PR8 virus strain has a molecular weight of 82,400 (Fields and Winter, 1982). There is evidence both from other sequence determinations (Bishop *et al.*, 1982 b; Robertson and Roditi, in preparation) and from two dimensional non-equilibrium pH gradient-SDS-polyacrylamide gel electrophoresis (Horisberger, 1980; Horisberger and de Staritzky, 1981) that the product of RNA segment 3 is an acidic protein in all strains of influenza A and B viruses so far examined.

Little is known regarding the function of PA. It is present, with PB1 and PB2, in transcriptive complexes isolated from infected cells which also include NP (Inglis *et al.*, 1976); more recently, direct isolation of an active transcriptive complex from PR8 virions which contained only PB1, PB2 and PA has been reported (Kawakami and Ishihama, 1982). Pulse-chase analysis of fowl plague virus-infected MDCK cells suggests, however, that PA differs from PB1 and PB2 in that whilst the latter two polypeptides migrate rapidly into the cell nucleus after synthesis (along with NP and NS_1), PA remains largely (if not entirely) in the cytoplasm (Briedis *et al.*, 1981; Mahy *et al.*, 1981 b).

Table 6. *Phenotype at the restrictive temperature of influenza A mutants having a* ts *lesion in RNA segment 2*

Mutant	RNA transcriptase		Virus-specific RNA			Virus-specific polypeptides		Comments
	Virion	Induced	Total[a]	A(+)[b] cRNA	A(−) cRNA and vRNA[b]	Early	Late	
C1	+							
C15	+			±	−	+	+	
G18		−	+					These mutants are defective in transport of the polymerase complex out of the nucleus
G90		±	+					
G93		±	+					
G236		−	−					
M131	−		−	+	−	−	−	
NYT5			±			−	−	
NYT15			−		−	−	−	
NYT101	±		−		−	−	−	
B315				±	−			

[a] Measured by ^{3}H-uridine labelling in presence of actinomycin D.
[b] Measured by hybridization to specific probes.

a) Virion Transcriptase Activity

In their survey of transcriptase activity of WSN influenza virus mutants, Mowshowitz and Ueda (1976) recorded no significant temperature lability for *ts* NYT53 virion transcriptase activity. Similarly, the Moscow segment 3 mutant *ts* M166 behaved like wild-type virus when analyzed for *in vitro* transcriptase activity at 36 °C and 42 °C (Ghendon *et al.*, 1973), and no temperature-sensitive defects were detected with mutants *ts* C45 or *ts* CmN4 when analyzed in the presence of either ApG or globin mRNA as *in vitro* primer (Nichol *et al.*, 1981). These reports would be consistent with lack of involvement of PA in the *in vitro* transcriptase reaction, but only four mutants have been studied so far, and further phenotypic analysis is necessary before firm conclusions are possible.

b) RNA-Dependent RNA Polymerase Activity of Infected Cell Extracts

Scholtissek and Bowles (1975) found that an RNA segment 3 mutant (*ts* G263) failed to induce RNA-dependent RNA polymerase activity in the cytoplasm of cells maintained at 40 °C throughout a 5.5 hour infection period. If the infection was at 33 °C for 4 hours, followed by a shift to 40 °C from 4.5–5 hours post-infection, polymerase activity was induced and remained at normal levels. It was concluded that *ts* 263-induced polymerase activity was stable at 40 °C, even if cycloheximide was present during the 1.5 hour incubation at the higher temperature (Scholtissek and Bowles, 1975). This supports the data so far obtained with the virion RNA transcriptase activities of mutants in this group.

c) Virus-Specific RNA Synthesis *in vivo*

Early studies in which RNA synthesis was measured by ^{3}H-uridine labelling in the presence of actinomycin D added at 3–4 hours post-infection have generally shown that segment 3 mutants are RNA negative; as discussed previously, this procedure probably detects only new vRNA synthesis. Ghendon *et al.* (1973) originally found that *ts* M166 was RNA negative using this procedure. In a later study, cells were incubated with ^{3}H-uridine from 2–4 hours post-infection in the absence of actinomycin D, and virus-specific cRNA was determined by hybridization. It was found that *ts* M166 had no defect in cRNA synthesis (Ghendon *et al.*, 1975).

Studies of WSN mutants in which ^{3}H-uridine incorporation was measured in the presence of actinomycin D showed that 7 group II (segment 3, Ritchey and Palese, 1977) mutants, *ts* NYT 4, 7, 52, 53, 57, 59, and 103 were RNA negative (Sugiura *et al.*, 1975). Hybridization analysis of *ts* NYT53 showed that cRNA synthesis was reduced at the higher temperature, but was multiplicity-dependent, since more cRNA synthesis was observed at high multiplicity of infection (in contrast to wild-type virus) (Krug *et al.*, 1975). When *ts* NYT53-infected cells were shifted to 39.5 °C after 4 hours at 33 °C, synthesis of cRNA continued at the same or a higher rate than was observed at 33 °C. It was included that the group II mutants are not defective in the transcriptase, but most likely are defective in synthesis of vRNA (Krug *et al.*, 1975). Direct measurement of both cRNA and vRNA synthesis in

ts NYT53-infected cells has recently confirmed that cRNA synthesis is largely unimpaired at the restrictive temperature whereas vRNA synthesis is negative with this mutant (Mowshowitz, 1981). It is interesting that another mutant in this group, *ts* NYT52, readily induces defective interfering virus formation at the permissive temperature: the other group II mutants tested did not share this property (Nayak *et al.*, 1978).

A fowl plague virus mutant, *ts* C45, is also vRNA negative (Mahy *et al.*, 1981 a). No defect in primary transcription was found with this mutant by comparing cells infected at 40.5 °C with 34 °C. However, accumulation of A(−) cRNA and vRNA were both less than 10% at the restrictive temperature, and shift experiments confirmed that very little vRNA was synthesized between 2 and 4 hours at the restrictive temperature, even if cells were allowed to accumulate A(−) cRNA at the permissive temperature during the first 2 hours of infection (Mahy *et al.*, 1981 a). One of the Giessen fowl plague virus mutants defective in RNA segment 3, *ts* G263, was also found negative for vRNA synthesis both at 40 °C throughout infection, and after shift-up to 40.5 °C following a 3.5 hour period at 33 °C (Scholtissek and Bowles, 1975). A more recent study using the Bethesda mutants also pointed to a role for the segment 3 product in vRNA synthesis. Thierry and Danos (1982) found that *ts* B454 produced A(−) cRNA normally at 39 °C, but failed to make significant amounts of vRNA. These studies all point to a direct role of PA in the synthesis of vRNA.

d) Virus-Specific Polypeptide Synthesis

Rather variable results have been reported for different RNA segment 3 mutants with respect to virus-induced polypeptide synthesis. Ghendon *et al.* (1973) found no defect in induced virus protein synthesis with *ts* M166 infection at 42 °C compared to 36 °C. Similarly, Sugiura *et al.* (1975) reported that syntheses of RNP, HA and NA were all similar to wild-type virus infection in cells infected at 39.5 °C with *ts* NYT 4, 7, 52, 53, 57, 59, and 103. However Almond (1978) found no virus-induced polypeptides in cells infected at the restrictive temperature with *ts* C SP10. Another Cambridge mutant, *ts* C45, induces at 40.5 °C the synthesis, at a low rate, of similar amounts of all the virus-specific proteins except NS_2 (which is encoded in a spliced mRNA) (Mahy *et al.*, 1981 a). This is similar to the pattern of polypeptide synthesis observed in wild-type virus-infected cells after release from a cycloheximide block to allow only primary transcription (Inglis and Mahy, 1979). It was suggested that *ts* C45 may be blocked in the replication cycle at the point immediately following primary transcription and so is unable to initiate synthesis of full-length template cRNA and subsequent amplification of vRNA (Mahy *et al.*, 1981 a).

e) Summary

The results obtained with RNA segment 3 mutants indicate that PA is involved in the synthesis of vRNA in infected cells (Table 7). Since this process is the least understood of the various steps in genome formation following infection, further study of *ts* mutants in this group should prove very rewarding. No defects in virion

Table 7. *Phenotype at the restriction temperature of influenza A mutants having a* ts *lesion in RNA segment 3*

Mutant	RNA transcriptase		Virus-specific RNA				Virus-specific polypeptides		Comments
	Virion	Induced	Total[a]	A(+) cRNA[b]	A(−) cRNA[b]	vRNA[b]	Early	Late	
C45	+			+	−	−	±	−	Protein synthesis blocked at primary transcription stage
CmN4	+								
C SP10							−	−	
G263		+				−			
M166	+		−	+	−			+	Protein synthesis assessed as accumulation of virus-specific RNP, HA and NA
NYT4			−					+	
NYT7			−					+	
NYT52			−					+	
NYT53	+		−	+	+	−		+	
NYT57			−					+	
NYT59			−					+	
NYT103			−					+	
B454					+	−			

[a] Measured by ^{3}H-uridine labelling in presence of actinomycin D.
[b] Measured by hybridization to specific probes.

transcriptase activity or in primary transcription have so far been recorded for segment 3 mutants; thus there is no evidence that PA plays a direct role in virus-specific mRNA synthesis.

5. *Mutants with* ts *Lesions in RNA Segment 4*

RNA segment 4 encodes the haemagglutinin polypeptide, which forms one of the trimer of subunits making up the mature haemagglutinin spike (Wilson *et al.*, 1981). In the PR8 strain influenza virus, the RNA segment is 1765 residues long; the mRNA synthesized from this segment encodes a precursor haemagglutinin molecule which is modified by glycosylation and cleavage to form the larger HA1 and the smaller HA2 subunit (Winter *et al.*, 1981 a). A number of *ts* mutants with lesions in RNA segment 4 have been described; where it has been studied, no defects in virus-specific RNA synthesis have been recorded with these mutants.

Sugiura *et al.* (1975) noted briefly that one of their group VI mutants, *ts* NYT104, was defective in haemagglutination, but this turned out to be too leaky for biochemical analysis. A detailed study of another mutant, *ts* NYT61, showed it to be a double mutant (groups V and VI) which had an unusually heat-labile haemagglutinating activity: segregation of single mutants *ts* NYT61 resulted in one, classified as group VI, which retained the haemagglutinin defect and was termed *ts* NYT61S (Ueda and Kilbourne, 1976). Cells infected with *ts* NYT61S at a multiplicity higher than 1 p.f.u. per cell developed cytopathic effects at the non-permissive temperature but neither haemagglutinating nor neuraminidase activity was detected in the culture fluid. However, neuraminidase activity, but not haemagglutinin, was found in extracts of cells kept at the non-permissive temperature during infection. Progeny were rescued from *ts* NYT61S-infected cells at 39.5 °C by mixed infection at the permissive temperature with a temperature-resistant H3N1 recombinant, and those which had the HA serotype of the *ts* parent were all temperature-sensitive, localizing the *ts* defect in the haemagglutinin gene (Ueda and Kilbourne, 1976).

Two defects were evident in *ts* NYT61S-infected cells. First, neither the HA nor its unglycosylated precursor (HA_o) could be detected in cells infected at the restrictive temperature and labelled with [^{14}C]-glutamine or [^{14}C]-fucose. Second, even though produced at the permissive temperature, the HA of *ts* NYT61S was unusually heat-labile, and appeared to be less susceptible to cleavage into HA_1 and HA_2 in the host MDBK cells than ts^+ virus (Ueda and Kilbourne, 1976). In a later paper it was reported that *ts* NYT61S did in fact synthesize HA at the non-permissive temperature, but in a partially glycosylated form which differed from mature ts^+ HA in gel mobility (Palese, 1978).

Four of the Giessen mutants had *ts* defects in RNA segment 4 and two of these, *ts* G227 and 79, were defective in HA activity (Scholtissek and Bowles, 1975). Detailed analysis of *ts* G227 showed that the HA was synthesized at the restrictive temperature, and incorporated into the rough endoplasmic reticulum, but was not transported to the smooth endoplasmic reticulum or to the cell surface (Lohmeyer and Klenk, 1979). Presumably the HA formed at 40 °C lacks a structural feature necessary to promote transport to the cell surface. This *ts* defect in migration

apparently prevents complete glycosylation and proteolytic cleavage of the mutant HA; mannose and glucosamine are incorporated, but galactose and fucose are only attached at the permissive temperature. As with *ts* NYT61S, synthesis of neuraminidase in *ts* G227-infected cells is normal at the restrictive temperature, indicating that there is little interdependence in the processing of the two envelope glycoproteins (Lohmeyer and Klenk, 1979).

Three of the Cambridge mutants were found to have *ts* lesions in RNA segment 4, and one of these, *ts* C US4 was temperature-sensitive for haemagglutination (Almond *et al.*, 1979). HA tests with this mutant can be carried out at 34 °C, but at 40.5 °C haemagglutination of erythrocytes does not occur (Almond, 1978). Virus-erythrocyte complexes formed at 34 °C dissociate if placed at 40.5 °C. Another of these mutants, *ts* C46, fails to synthesize detectable amounts of the HA polypeptide in cells infected at 40.5 °C; neither the mature HA nor the unglycosylated precursor molecule is formed (J. W. McCauley and P. Colloby, unpublished observations). It is not known at which level the block in HA synthesis occurs with this mutant.

6. *Mutants with* ts *Lesions in the RNA Segments Encoding the Nucleoprotein*

The fifth largest RNA segment encodes the nucleoprotein (NP). In the PR8 strain, the NP is a basic, arginine-rich molecule comprised of 498 amino-acids (Winter and Fields, 1981). It has generally been assumed that transcription of virus RNA occurs in complexes consisting of one or more P proteins as well as NP, and such complexes have been isolated from infected cells and shown to be transcriptionally active (Schwartz and Scholtissek, 1973; Compans and Caliguiri, 1973; Caliguiri and Compans, 1974; Inglis *et al.*, 1976; Rochovansky, 1976). However, a recent report described the isolation of an active RNA transcriptase complex containing three P proteins only, the NP having been removed by two cycles of Cs_2SO_4 centrifugation (Kawakami and Ishihama, 1982). A number of mutants with *ts* lesions in the NP gene have been found defective in the later stages of the virus replication cycle, including vRNA synthesis and virus maturation.

a) RNA Transcriptase Activity

In their study of *in vitro* transcriptase activity of *ts* NYT mutant virions, Mowshowitz and Ueda (1976) examined one group V mutant, *ts* NYT56 and found no evidence of temperature sensitivity, although the range studied was below the restrictive temperature. Similarly, during assays of *ts* C34 and *ts* CUS1 in the presence of either ApG or globin mRNA added as *in vitro* primer molecules, RNA transcriptase activity was comparable to that of ts^+ virus when assayed at 40.5 °C (Mahy *et al.*, 1981 a; Nichol *et al.*, 1981).

The induced RNA-dependent RNA polymerase activity of four Giessen group VI *ts* mutants was reported by Scholtissek and Bowles (1975). Three, *ts* G19, *ts* G71 and *ts* G283, induced at 40.5 °C about 50—70% of the polymerase activity found in cells kept at 33 °C. It should be noted that the ts^+ virus in these experiments induced 30—50% higher polymerase activity at 40 °C compared to 33 °C, but these two mutants were nevertheless regarded as RNA polymerase positive. The other NP

mutant, *ts* G81, appeared to be polymerase negative, since less than 5% of the activity induced at 33 °C was observed in cells infected at 40.5 °C. All assays in these experiments were at 32 °C, so it was not established in these experiments whether the induced RNA-dependent RNA polymerase activity, once formed, was temperature-sensitive (Scholtissek and Bowles, 1975).

b) Virus-Specific RNA Synthesis *in vivo*

Sugiura *et al.* (1975) reported that *ts* NYT56 was negative for RNA synthesis *in vivo,* although another NP gene mutant, *ts* NYT12, incorporated about 30% of the amount of ^{3}H-uridine in presence of actinomycin D as compared to ts^+ virus. It was considered possible that *ts* NYT12 was a leaky mutant, and this was also concluded in a later study by Krug *et al.* (1975). The latter authors found that *ts* NYT56 had no defect in cRNA synthesis, and continued to synthesis cRNA after shift-up to the restrictive temperature at 5 hours post-infection. Although not measured directly in this study, it was concluded that synthesis of vRNA was defective in *ts* NYT56-infected cells (Krug *et al.,* 1975).

One of the Giessen group VI mutants, *ts* G19, appeared to have no defect in RNA synthesis, whereas another, *ts* G81, was negative for vRNA synthesis (Scholtissek and Bowles, 1975). Neither cRNA nor vRNA were synthesized in *ts* G81-infected cells kept continuously at 40.5 °C. However, if cells were maintained at 33 °C for 4 hours after infection and then shifted up to 40 °C for a further 2 hours, very little vRNA was synthesized despite the presence of normal amounts of cRNA (Scholtissek, 1978). The defect in *ts* G19, which was not apparently defective in virus-specific RNA synthesis, was considered to be in some stage of virus maturation (Scholtissek, 1979).

The only Cambridge NP mutant which has been studied, *ts* C US1, synthesized no A(−) cRNA or vRNA in cells maintained at 40.5 °C throughout infection. However, if cells were maintained at 34 °C for 2 hours and then shifted up to 40.5 °C, A(−) cRNA accumulated normally, but vRNA accumulated to only 40% of the concentration reached in cells kept at 34 °C throughout (Mahy *et al.,* 1981 a). The phenotype of *ts* C US1 is therefore similar to that of *ts* G81.

Two NP mutants of the Hong Kong virus strain have been analyzed using single-stranded DNA probes prepared by M13 cloning. As with the Giessen mutants, the two Bethesda mutants had quite different phenotypes: *ts* B463 had no detectable *ts* defects in virus-specific RNA synthesis, whereas *ts* B2C synthesized reduced amounts of A(+) cRNA and no vRNA at the restrictive temperature (Thierry and Danos, 1982).

c) Virus-Specific Polypeptide Synthesis

Reduced production of RNP, HA and NA was found in cells infected by *ts* NYT12 and *ts* NYT56 (Sugiura *et al.,* 1975), but further protein studies with these mutants have not been reported. Scholtissek and Bowles (1975) analyzed the virus-induced polypeptide pattern in cells infected with *ts* G19; most polypeptides were synthesized normally at 40 °C, but there appeared to be a reduction in synthesis of M and/or NS_1 protein(s). The gel slicing technique used in these experiments gave

Table 8. *Phenotype at the restriction temperature of influenza A mutants having a* ts *lesion in the RNA segment encoding the nucleoprotein*

Mutant	RNA transcriptase		Virus-specific RNA				Virus-specific polypeptides		Comments
	Virion	Induced	Total[a]	A(+) cRNA[b]	A(−) cRNA[b]	vRNA[b]	Early	Late	
C34	+				+	±	+	+	
CUS1	+								
G19		±		+	+	+	+	±	Reduced synthesis of M/NS_1
G71		±							
G81		−		+	+	±			
G283		±							
NYT12			±					±	Reduced synthesis of RNP, HA and NA
NYT56			−	+	+	−		±	
B2C				±		−			
B463				+	+	+			

[a] Measured by ^{3}H-uridine labelling in presence of actinomycin D.
[b] Measured by hybridization to specific probes.

poor resolution in this region of the polyacrylamide gel. The pattern of virus-induced polypeptides in cells infected with *ts* C US1 was similar at the restrictive and non-restrictive temperatures (Mahy *et al.*, 1981 a).

d) Summary

Relatively few mutants with *ts* lesions in the NP gene have been subjected to phenotypic analysis (Table 8). Those that have been analyzed fall into one of two classes, a group in which vRNA synthesis is temperature-sensitive, and another in which virus-specific RNA and proteins are apparently made normally. In the latter case, it is assumed that a *ts* defect exists in the virus maturation process. For example, NP is likely to be required to interact both with the RNA segments and with the matrix protein during assembly. Such a dual role of NP in virion RNA synthesis and assembly would be consistent with genetic evidence that intracistronic complementation occurs with this segment (Thierry *et al.*, 1980; Thierry and Spring, 1981; Shimizu *et al.*, 1982 b). There is no evidence from *ts* mutant analyses made so far that NP plays a role in mRNA synthesis, and this is consistent with the recent isolation of transcriptionally active complexes which contain only P proteins (Kawakami and Ishihama, 1982).

7. *Mutants with* ts *Lesions in the RNA Segment Encoding the Neuraminidase*

A number of neuraminidase-deficient mutants have been isolated, and their study has contributed substantially to knowledge of the role of this protein in virus replication. Scholtissek and Bowles (1975) noted that three mutants (*ts* G113, 210 and 288) belonging to recombination group IV produced at 40 °C HA which gave only incomplete haemagglutination in a standard test, and no detectable NA activity. When mutant-infected cells were incubated at 33 °C, neuraminidase activities were 1% (*ts* G113) or 10% (*ts* G210 and *ts* G288) of that found in ts^+ virus-infected cells. Virus release was not affected at 40 °C, and the *ts* mutant virus yield (titrated at 33 °C) was similar to that obtained at 33 °C in a single-cycle experiment. However, plaque formation by these three mutants was abnormal at 40 °C, and only tiny plaques developed during prolonged incubation (Scholtissek and Bowles, 1975).

Two WSN strain neuraminidase mutants (*ts* NYT3 and 11) studied by Palese *et al.* (1974) displayed a somewhat different phenotype. Neuraminidase activity of these mutants was thermolabile, even in virus produced at 33 °C. The most interesting results were obtained when mutant virus-infected cells were examined by electron microscopy (Palese *et al.*, 1974). At 39.5 °C, morphologically intact virus particles were observed in these cells, but large aggregates of virions accumulated near the cell surface. The envelopes of these aggregated viruses contained neuraminic acid as revealed by colloidal iron staining, and it was concluded that this neuraminic acid serves as a receptor for the haemagglutinin on other virions, thus aggregating the particles and preventing their release from the cell surface. Treatment with bacterial neuraminidase both destroyed the aggregates formed at

the restrictive temperature and restored HA activity. These results suggested that the virus neuraminidase is essential for replication not because it is required for virus release from the cell, but in order to prevent particle aggregation by the haemagglutinin (Palese *et al.*, 1974).

A recent study revealed a somewhat different phenotype for *ts* M5, which has a lesion in RNA segment 6 encoding the neuraminidase of the fowl plague virus Weybridge strain (Ghendon *et al.*, 1979; 1983 a). Synthesis of all species of virus-specific RNA was detected at the non-permissive temperature in cells infected with *ts* M5, and all virus-specific proteins were synthesized. However comparisons with ts^+ virus-infected cells showed that the regulation of both RNA and protein synthesis was altered in *ts* M5-infected cells at the non-permissive temperature. The synthesis of RNA segments 1, 2, 3, 4, 6, and 7 (both cRNA and vRNA), as well as of their corresponding polypeptides, was reduced during the late phase of the growth cycle (from 3 hours post-infection) and virus morphogenesis was impaired. Both the neuraminidase and the haemagglutinin synthesized under non-permissive conditions functioned normally and migrated from the rough endoplasmic reticulum into plasma membranes, although haemagglutinin cleavage was reduced (Ghendon *et al.*, 1983 a). How a mutation in the neuraminidase gene which does not affect enzymatic activity can prevent virion formation is not clear, and it is even more difficult to understand how neuraminidase might contribute to the regulation of RNA synthesis and transcription. Nevertheless, these results suggest that the role of the neuraminidase during virus replication may not merely be to prevent virus aggregate formation.

8. *Mutants with* ts *Lesions in RNA Segment 7*

RNA segment 7 encodes the matrix (M) polypeptide (mol. wt. 28,000), the most abundant protein species in the virion, as well as one or possibly two further polypeptides of unknown function. Sequence analysis of this segment from the PR8 virus strain revealed two open reading frames (Allen *et al.*, 1980; Winter and Fields, 1980) and these have also been found in RNA segment 7 of fowl plague virus (McCauley *et al.*, 1982) and the A/Udorn/72 (H3N2) strain (Lamb and Lai, 1981 a). However, analysis of segment 7-specific RNA in virus-infected cells showed that there were not two but three mRNAs derived from this segment by a splicing process (Inglis, 1981; Inglis and Brown, 1981; Lamb and Lai, 1981 b; Lamb *et al.*, 1981). In addition to the M protein, a second protein (M2) of unknown function has been found in a variety of influenza virus-infected cells (Lamb and Choppin, 1981), but the product of the third mRNA transcribed from segment 7 (? M3) has not been identified (see Lamb, this volume).

The M protein sequence is highly conserved in those influenza virus strains examined so far, and this, coupled with the small target size of this segment and the existence of an overlapping reading frame, may be responsible for the fact that only three mutants with *ts* lesions in segment 7 have been described. Sugiura *et al.* (1975) reported two group VII mutants, *ts* NYT51 and 62, subsequently found to have lesions in RNA segment 7 (Ritchey and Palese, 1977). No defects in RNA synthesis or in the production of RNP, HA or NA were recorded during infection at non-

permissive temperature by these mutants, and induced cytopathic effects were equivalent to those seen in wild-type virus infection (Sugiura *et al.*, 1975). All virus-specific polypeptides were synthesized at the non-permissive temperature in mutant-infected cells (unpublished observations of K. Tobita and P. Palese, reported in Ritchey and Palese, 1977). Further phenotypic analyses of *ts* NYT51 and *ts* NYT62 have not, however, been reported.

Recently, Ghendon *et al.* (1982 b) have reported phenotypic studies on *ts* M303/1, which was shown to have a *ts* lesion in RNA segment 7 by crossing with a different fowl plague virus strain, Rostock, in which all eight genome segments differ in mobility (Markushin *et al.*, 1981). Virions of *ts* M303/1 appeared to be deficient in haemagglutinating activity even when grown in eggs at the permissive temperature, and had an ID_{50}/HA ratio which was 20-fold higher than that of ts^+ virus. Synthesis of all species of RNA and polypeptides occurred in *ts* M303/1-infected cells at the non-permissive temperature, but syntheses of functionally active haemagglutinin and neuraminidase were not observed. In particular, cleavage of HA into HA_1 and HA_2 did not occur at the restrictive temperature, and was reduced compared to ts^+ virus even at the permissive temperature. Studies of recombinant viruses between *ts* M303/1 and *ts* C46 showed that the HA of *ts* M303/1 contained no *ts* lesion; similar crosses with A/England/42/72 confirmed that the neuraminidase of *ts* M303/1 was not defective either. It was concluded that the impairment in HA cleavage of this mutant is due neither to an altered HA molecule, nor to a change in the neuraminidase which might in turn affect haemagglutinin cleavage (Schulman and Palese, 1977). Most likely, some conformational change in the M protein affects the maturation of HA and neuraminidase in this mutant, although the possibility that the *ts* defect is expressed through a change in the M_2 or the M_3 protein induced by *ts* M303/1 cannot be ruled out (Ghendon *et al.*, 1982 b). It would seen desirable to isolate further segment 7 *ts* mutants, which should prove valuable in exploring the role of the small RNA segment 7 protein(s) during influenza virus replication.

9. *Mutants with* ts *Lesions in RNA Segment 8*

The smallest genome RNA segment of influenza viruses encodes two non-structural proteins, NS_1 and NS_2, which are synthesized from spliced mRNA's encoded in overlapping reading frames (Inglis *et al.*, 1979; Lamb and Choppin, 1979; Inglis *et al.*, 1980; Lamb and Lai, 1980). NS_1 (mol. wt. 26,000) is an early virus-specific protein which rapidly migrates into the cell nucleus after synthesis in the cytoplasm (Dimmock, 1969; Briedis *et al.*, 1981); late in the virus growth cycle, NS_1 forms cytoplasmic inclusions with some influenza virus strains (Morrongiello and Dales, 1977; Shaw and Compans, 1978) but not others, *e.g.* fowl plague virus (Petri *et al.*, 1982). The inclusions appear to consist of NS_1 complexed to RNA (Yoshida *et al.*, 1981); that they are not seen in some virus infections may reflect variations in charge of the NS_1 protein, which is basic in A/Udorn virus-infected cells, but acidic in fowl plague virus-infected cells (Petri *et al.*, 1982). NS_2 (mol. wt. 14,000), accumulates in infected cells later than NS_1, and remains in the cytoplasm (Briedis *et al.*, 1981; Mahy *et al.*, 1980).

Only three *ts* mutants with defects in RNA segment 8 have so far been described in any detail, and these were derived from fowl plague virus Rostock strain. Presumably, as with segment 7, the isolation of mutants is restricted by small target size, conserved sequence, and the presence of overlapping reading frames. However Shimizu *et al.* (1982 b) have recently found four further segment 8 mutants in their collection of 136 A/Udorn mutants. Phenotypic analysis of these mutants has not so far been reported.

Almond (1978) isolated two *ts* mutants with lesions in RNA segment 8. One, *ts* C47, induces the synthesis of NS1 polypeptide with faster than normal mobility (Almond *et al.*, 1977), and an estimated mol. wt. of 23,000 (Robertson *et al.*, 1983). This is a stable phenotype which is not temperature-dependent; recent sequence analysis of *ts* C47 segment 8 has revealed a nonsense mutation at nucleotide 634 at which a C residue is replaced by a U residue in the genome RNA (Robertson *et al.*, 1983). This G to A change in the transcript results in the conversion of a UGG (trp) codon in the mRNA sequence into a UGA (opal) nonsense codon. However, the single base change lies within the overlapping reading frame, and in that frame converts a GGA (gly) codon to a GAA (glu) codon in the NS2 mRNA. It is therefore not clear whether the phenotypic changes observed with *ts* C47 result from alterations in NS1, NS2 or both polypeptides. It is interesting that the loss of 28 amino acids from the carboxy terminus of the NS1 induced by *ts* C47 does not apparently affect growth of this mutant at the permissive temperature.

Phenotypic characterization of *ts* C47 showed that the virion transcriptase activity at the restrictive temperature was similar to that of ts^+ virus, and responded normally to priming by globin mRNA or ApG (Nichol *et al.*, 1981; Wolstenholme *et al.*, 1980). Synthesis of A(+) cRNA also occurred normally in *ts* C47-infected cells at 40.5 °C and A(−) cRNA increased as in ts^+ virus-infected cells up till 2.5 hours post-infection. However vRNA synthesis in *ts* C47-infected cells at 40.5 °C was reduced to only 6% of that found at 34 °C. Temperature-shift experiments with this mutant showed that shifting up to 40.5 °C at any time up to 3 hours post-infection blocked virus replication (Wolstenholme *et al.*, 1980).

In addition to the defect in vRNA synthesis, alterations in the regulation of virus-specific polypeptide synthesis were found in *ts* C47-infected cells. The switch from early to late protein synthesis (Inglis and Mahy, 1979) occurred normally at 34 °C but did not occur at 40.5 °C so that both HA and M syntheses were greatly reduced. Although NS2 was only detected at the permissive temperature in these experiments (Wolstenholme *et al.*, 1980), more recently NS2 has also been detected in *ts* C47-infected cells at 40.5 °C, although the relative amount of mRNA for this protein is reduced (D. B. Smith and S. C. Inglis, unpublished observations).

Similar studies carried out on *ts* C mN3 showed that this mutant induces a similar pattern of virus-specific RNA synthesis in infected cells to *ts* C47 (Wolstenholme, *et al.*, 1980). In particular, no vRNA is synthesized at the restrictive temperature, although A(−) cRNA and A(+) cRNA accumulate at similar rates at the upper and lower temperatures. The presence of cRNA in the virtual absence of vRNA with these mutants suggests that most cRNA may be transcribed from the infecting virion RNA.

Polypeptide synthesis in *ts* mN3-infected cells was confined, at 40.5 °C, to an early pattern, as with *ts* C47. However the NS_2 protein was not detected at either upper or lower temperature in these experiments (Wolstenholme *et al.*, 1980). More recently, the mRNA encoding NS_2 has been identified in *ts* mN3-infected cells at the permissive temperature by blot hybridization using specific cloned DNA fragments derived from fowl plague virus RNA segment 8 (D. B. Smith and S. C. Inglis, unpublished observations). However, the concentration of NS_2 mRNA (in cells at the permissive temperature) was more than tenfold lower in mutant-infected cells, compared to ts^+ virus-infected cells, and the NS_2 polypeptide has not so far been found in *ts* CmN3-infected cells at either temperature. It is of course possible that NS_2 induced by *ts* CmN3 has an altered mobility, and experiments to explore this are in progress.

A third RNA segment 8 mutant, *ts* G412, appears to have a late defect in virus replication (Koennecke *et al.*, 1981). This mutant was derived by 5-FU mutagenesis of a fowl plague virus Rostock-virus N recombinant (90/N3) which carries RNA segments 2 and 8 from virus N and all other segments from fowl plague virus. All virus-specific RNA species are synthesized normally at the non-permissive temperature, and RNA-dependent RNA polymerase activity is not reduced in *ts* G412 mutant-infected compared to parental virus-infected cells. The only clear defect, other than failure to produce infectious particles, is that the switch from early to late protein synthesis does not occur (Koennecke *et al.*, 1981). In this respect, the phenotype of *ts* G412 resembles *ts* C47 and *ts* CmN3.

Scholtissek and Spring (1981 a and b) have reported two additional *ts* mutants derived by mutagenesis of virus 19N, a possible partial heterozygote virus (Scholtissek *et al.*, 1978 b). One (*ts* G451) was a double mutant with lesions in segments 5 and 8; the other (*ts* G526) could be rescued by recombination with *ts* mutants of all Giessen groups I to VI, but did not recombine with *ts* G412, so it was assumed to be a single mutant with a lesion in segment 8. However the possibility that segment 7 might also be defective could not be tested. No detailed phenotypic analyses of these mutants have been reported, but two interesting observations were made in these experiments. First, replacement of RNA segments 2, or 1 and 3 together, by segments from other influenza virus strains restored normal growth properties to these mutants, although the segment 8 *ts* lesion was retained. This phenomenon, which was termed "suppressor recombination" (Scholtissek and Spring, 1981 a, b) suggests that the products of the polymerase segments co-operate with the non-structural protein(s) during replication, and is consistent with the vRNA-negative phenotype observed for *ts* C47 and *ts* CmN3 (Wolstenholme *et al.*, 1980).

The second feature of interest was the finding that in at least one case the suppression of the *ts* phenotype of a *ts* G526 recombinant occurred in chick embryo fibroblast cells but not in MDCK cells (Scholtissek and Spring, 1981 b). This suggests that a host factor may also be required to co-operate with the gene product of segment 8 and/or with the product of segments 1 to 3 during replication. Host-dependent *ts* mutants of A/Udorn virus have also been described by Shimizu *et al.* (1982 a) and will be considered in more detail in section V.

D. Temperature-Sensitive Mutant-Derived Vaccines

It was suggested a number of years ago that *ts* mutants of influenza viruses might be attenuated, and therefore worthy of investigation in the construction of live vaccines (Mackenzie, 1969). The lower temperature (32—34 °C) of the upper respiratory tract should allow multiplication of *ts* viruses which would be restricted in replication in the lungs and lower respiratory tract. A considerable effort in evaluating this approach has been made by Chanock, Murphy and co-workers, and these investigations have recently been well reviewed (Murphy *et al.*, 1980 a). The two most widely studied *ts* donor viruses which have emerged are the *ts*-1[E] and the *ts*-1A2 recombinants. The original *ts*1 parent was derived by 5-fluorouracil mutagenesis from an Asian (H2N2) strain, A/Great Lakes/65 (Mills and Chanock, 1971) and recombined with an A/Hong Kong/68 (H3N2) strain to yield the HK/68-*ts*-1[E] mutant, which was found to have *ts* lesions on RNA segments, encoding the PB2 and NP proteins (Palese and Ritchey, 1977). A second clone derived from the recombination of *ts*1 with the wild-type Hong Kong strain, designated *ts*-1[A], was recombined with the A/Udorn/72 (H3N2) strain to yield a clone (189) containing a single *ts* lesion in RNA segment 1, and this was recombined with a Hong Kong-Bethesda group 5 mutant (*ts* B315) having a *ts* lesion in RNA segment 2 to yield the A/Udorn/72-*ts*-1A2 recombinant, defective in both the PB2 and PB1 proteins (Massicot *et al.*, 1980 a).

Comparisons have been made both in experimental animals and in volunteers of the degree of attentuation of these two donor viruses by comparing recombinants of the A/Victoria/75 strain, namely A/Vic/75-*ts*-1[E] and A/Vic/75-*ts*-1A2 (Murphy *et al.*, 1980 a). The *ts*-1A2-bearing recombinants appeared to be somewhat more restricted in virulence than the *ts*1[E] recombinants, and to be genetically stable. However, extension of these studies to an A/Alaska/77-*ts*-1A2 recombinant which was used to vaccinate a seronegative child resulted in reversion to a wild-type phenotype, associated with multiple genetic changes including an extragenic suppressor and other mutations (Tolpin *et al.*, 1981; Murphy *et al.*, 1980 b). Although the patient did not develop symptoms, the revertant (designated FV1319) grew to high titre *in vivo*, and had comparable virulence to the A/Alaska/77 wild-type virus when administered intranasally to seronegative human volunteers (Tolpin *et al.*, 1981).

The genetic instability of these condidate vaccine donor strains illustrates the remarkable capacity of the influenza virus genome to undergo variation, even in those genes encoding internal proteins which are relatively free from antigenic selection pressure. If *ts* mutations are to form the basis of future attenuated influenza virus vaccines, some way will have to be found to overcome this problem.

IV. Cold-Adapted Mutants

One approach to the development of attenuated influenza virus strains suitable for vaccination of a human population is to adapt the virus to grow optimally at a low temperature, so that growth is restricted at 37 °C. The cold-adapted *(ca)* property may then be transferred to suitable candidate vaccine strains by gene reassortment.

Suitable donor cold-adapted virus strains have been developed in the U.S.S.R. from the A/Leningrad/9/46 (H1N1) and the A/Leningrad/134/57 (H2N2) viruses by Alexandrova and co-workers (Alexandrova and Smorodintsev, 1965; Alexandrova *et al.*, 1979), and in the U.S.A. from A/Ann Arbor/6/60 virus by Maassab and co-workers (Maassab, 1967; Maassab *et al.*, 1978, 1981; Odagiri *et al.*, 1982). These viruses were all found to be temperature-sensitive, and either fail to plaque at 39 °C, or grow poorly in eggs incubated at 40 °C. The use of these viruses for preparation of attenuated recombinant vaccine strains and in clinical trials has recently been reviewed in detail (Kendal *et al.*, 1981) and will not be considered here. However, the genetic basis of the *ca* phenotype has recently been explored with both the Leningrad and Ann Arbor viruses, and these investigations will be briefly reviewed.

The *ts* lesions in the Leningrad *ca* viruses were located by recombination with fowl plague virus mutants isolated in Moscow or Cambridge. Failure of recombination with specific mutants indicated lesions in the A/Leningrad/9/46 *ca* virus (passaged 37 times at 25 °C) at segments 1, 2, and 7, and in the A/Leningrad/134/57 virus (passaged 47 times at 25 °C) at segments 1, 2, 5, 7, and 8 (Kendal *et al.*, 1981).

The A/Ann Arbor/6/60 *ca* mutants were similarly analyzed by recombination experiments with *ts* NYT mutants of the WSN strain by Cox *et al.* (1981 a). Recombinations occurred with all groups studied except group III, with a lesion in RNA segment 2, implying that this segment carried the *ts* lesion in the *ca* virus. However this contradicted an earlier result in which similar recombination experiments with the Bethesda *ts* mutants of the Hong Kong strain had located a *ts* lesion on the Ann Arbor *ca* mutant RNA segment 1 (Spring *et al.*, 1977 a, b). To resolve this discrepancy, a series of recombinants was prepared at 33 °C between the *ca* mutant and either certain *ts* NYT mutants or an H3N2 ts^+ virus, A/Ann Arbor/9/73. Clones which were either ts^+ or *ts* were analyzed to determine their RNA segment derivation, but no single RNA segment correlated with the *ts* property, and every segment of the *ca* mutant A/Ann Arbor/6/60 virus was present in one or more non-*ts* recombinant viruses (Cox *et al.*, 1981 a, b). The only correlation with the *ts* property was that *ca* mutant RNA segments 1 and 7 were always present in *ts* recombinant viruses. It was therefore concluded that genetic synergism between the matrix and PB2 proteins somehow confers the *ts* property.

Aside from the *ts* property, biochemical studies of the A/Ann Arbor/6/60 *ca* virus show that every segment contains mutations when compared to the *wt* virus (Cox *et al.*, 1981 b). However, phenotypic analysis of the *ca* virus during infection of MDCK or chick embryo fibroblast cells at 39 °C revealed no defect compared to wild-type virus except for a slower migration rate of the P_2 protein (Cox *et al.*, 1981 b). It was concluded that the *ca (ts)* mutation affects a late replication function (Kendal *et al.*, 1981).

V. Host Range Mutants

The replication of influenza virus, either in cells in culture or in the animal kingdom as a whole, is characterized by a narrow range of permissive host cell species (Barry and Mahy, 1979; Pereira, 1982; Webster *et al.*, 1982). The factors

governing permissiveness of a given host cell type are poorly understood. Some cells are totally non-permissive, but in many cases influenza virus undergoes an abortive replication cycle which may be blocked at any stage up to the final maturation process (Caliguiri and Holmes, 1979). However, the end result of all such interactions is characterized by failure of the virus to form plaques on the restrictive host cell monolayer, and this property has been used to define host range *(hr)* mutants of influenza virus. Such mutants may be conditional (temperature-dependent, *td*) or non-conditional, and both types have been described; in addition, the ability of influenza viruses to undergo replacement of RNA segments by reassortment has led to the isolation of *hr* recombinant viruses.

A. Non-Conditional Host Range Mutants

Fowl plague virus does not normally form plaques on BHK or L cells, in which the replication cycle is abortive. Zavada (1969) adapted the Dobson strain of fowl plague virus (A/FPV/Dutch/27, H7N7) to plaque on BHK cells. The original Dobson strain was passaged five times and plaque purified three times on chick embryo fibroblast cells and designated Dobson wild-type virus. The *hr* mutant was derived by alternate growth on chick embryo fibroblast and BHK cells for six passages, followed by seven consecutive passages and one plaque purification in BHK cells. The resulting virus clone retained the ability to grow to high titres in eggs (J. Zavada, personal communication recorded in Almond, 1978).

The Dobson *hr* mutant was compared to Dobson wild-type virus by Israel *et al.* (1975), who noted a difference in migrational properties of the HA of the two strains on polyacrylamide gel electrophoresis, and this was independently confirmed in another report (Almond and Barry, 1978). However, the *hr* property of the Dobson strain appears to lie not in RNA segment 4 but in RNA segment 1. Almond (1978) obtained stocks of Dobson wild-type and the *hr* mutant cloned by Zavada from A. Israel (Paris), then plaque-purified the *hr* mutant three times in chick embryo fibroblast cells at 40.5 °C, and termed the resulting clone 4H. When this clone was compared to the original Dobson wild-type virus, both viruses were found to form plaques on BHK cells, although only the *hr* mutant would form plaques on L cells. Almond *et al.* (1977) prepared a large number of recombinants between fowl plague viruses Rostock (which did not plaque on BHK or L cells) and Dobson *hr* (clone 4H). The ability of certain of these recombinants to form plaques on either BHK cells or L cells correlated only with the presence of RNA segment 1 of 4H virus, which encodes PB2 (Almond, 1977; Almond and Barry, 1978). The possibility that HA cleavage might also be involved in the *hr* property was discounted on the basis of pulse-chase analysis of proteins formed in infected cells, which showed that HA cleavage occurred equally with the Rostock as with the Dobson strain in each cell type (Almond and Barry, 1978).

Support for the involvement of RNA segment 1 in host range was provided by an investigation of the ability of the Giessen *ts* mutants of fowl plague virus Rostock to form plaques on MDCK cells at the permissive temperature (33 °C) (Scholtissek and Murphy, 1978). All mutants formed plaques on chick embryo fibroblast cells,

but four RNA segment 1 mutants (*ts* G3, 115, 117, and 132) and two RNA segment 2 mutants (*ts* G90 and 93) failed to form plaques on MDCK cells. However, mutants of both these groups were genetically competent in MDCK cells, since *ts* G3 and *ts* G90 were able to complement each other and to form plaques in cells co-infected by the two mutants. It is not clear whether the *ts* lesion is responsible for the *hr* restriction in these mutants, but the properties were co-variant since revertants selected for loss of the *ts* lesion regained the ability to plaque in MDCK cells (Scholtissek and Murphy, 1978). Tuckova and co-workers (Tuckova *et al.*, 1971; Tuckova and Vonka, 1973; Anisimova *et al.*, 1977) have shown that influenza virus A/NWS-D will multiply in a human diploid fibroblast (LEP) or a hamster embryo fibroblast (HEF) cell line, but a closely related virus, A/WS-MK will not grow in either cell strain. A host range recombinant virus (r_{12}) isolated by crossing the parent strains had derived seven RNA segments from the WS parent and only segment 6 from the NWS parent (Ghendon *et al.*, 1979 b). This recombinant grows in HEF cells, apparently because haemagglutinin cleavage occurs in contrast to the WS parent, so that some defect in the neuraminidase of WS virus is responsible for the host restriction. However r_{12} virus does not multiply in LEP cells, and another of the NWS genes must confer the ability to grow in the human cells (Ghendon *et al.*, 1979 b).

The neuraminidase gene was also implicated in the ability of A/WSN virus to plaque in bovine kidney (MDBK) cells. Recombinants were prepared between A/WSN and either A/GM1 or A/Hong Kong viruses which do not plaque in MDBK cells. Plaques could be produced on MDBK cells by HK-WSN or FMI-WSN viruses carrying only the neuraminidase gene of WSN virus, whereas a WSN-HK virus recombinant containing only the HK neuraminidase failed to plaque (Schulman and Palese, 1977).

From both these studies it appeared that the function of the neuraminidase is to promote HA cleavage, necessary for plaque formation. However Scholtissek *et al.* (1978 a) described a fowl plague virus-A/Hong Kong recombinant (81/H0) which carries the HA gene of fowl plague virus but is unable to plaque on chick embryo fibroblast cells, although it can plaque on MDCK cells. This recombinant was derived by mixed infection with *ts* G81 or *ts* G19 and the Hong Kong strain in MDCK cells at the mutant restrictive temperature (39 °C) and had acquired RNA segments 1, 2, 5, 6, and 8 from Hong Kong, and segments 3, 4, and 7 from fowl plague virus. The ability to plaque in chick embryo fibroblast cells correlated with pathogenicity for chickens, and so recombinant 81/H0 was apathogenic (Scholtissek *et al.*, 1978 a). The correlation between gene constellation and pathogenicity is complex and outside the scope of this review; an authoritative account of factors involved in pathogenicity may be found in Rott (1979) and chapter 9, of this volume.

B. Conditional Host Range Mutants

Temperature-dependent host range *(td-hr)* mutants which have an extended host range only at the permissive temperature provide a useful approach to identification of *hr* genes. Israel (1980 a) has characterized one such mutant *(td4)* which he

isolated by 5-fluorouracil mutagenesis of the Dobson *hr* mutant of fowl plague virus, and selected for its ability to plaque on L cells at 34 °C but not at 39.5 °C. This mutant was not *ts* for growth on chick embryo fibroblast cells. The *ts* defect was located on segment 1 of *td4*, since replacement of that segment with a fowl plague virus Rostock segment restored the ability to form plaques in L cells at 39.5 °C. Phenotypic analysis of this mutant suggested that vRNA synthesis was defective in L cells, but not in chick cells, at 39.5 °C (Israel, 1980 a).

Shimizu *et al.* (1981, 1982 a, 1983) have described 36 or their A/Udorn *ts* mutants which have a temperature-dependent host restriction. Four of these mutants were restricted in ability to plaque on rhesus monkey kidney (RMK) cells, but could plaque on MDCK cells at 40 °C. The other 32 mutants (designated type III) plaqued on RMK cells but were restricted in ability to plaque on MDCK cells at 40 °C. Sixteen of these 32 type III mutants displayed a reduction of more than 10,000-fold in their plaquing ability on MDCK cells at 40 °C. The locus of the *ts* lesion was determined for 15 of these mutants: 5 had a mutation in segment 2, 2 in segment 3, 4 in segment 6 (NP) and 4 in segment 8. It was concluded that RMK cells contain host factors which suppress the *ts* defect in these *td-hr* mutants (Shimizu *et al.*, 1983).

It is possible that a genetic interaction similar to that exhibited by *td-hr* mutants underlies the temperature sensitivity of viruses which are apathogenic in chickens. Rott *et al.* (1982) found that only pathogenic recombinant viruses are capable of multiplying normally at 41 °C, the body temperature of birds, whereas pathogenic and apathogenic viruses grow equally well at 37 °C. In each case where a pathogenic recombinant was obtained in crosses of fowl plague virus Rostock and A/Turkey/England/63 virus, RNA segments 1, 2, and 3 were all derived from one parent. Mixed derivation of segments 1 to 3 was associated with loss of pathogenicity for chickens as well as loss of ability to grow at 41 °C (Rott *et al.*, 1982).

C. Conclusions

The results obtained from analysis of host range mutants of influenza viruses point to a close co-operation between the P genes and some as yet undetermined host cell factor during virus replication. PB2, in particular, has been implicated as a determinant of host range in several studies. This protein is known to interact with capped host cell mRNAs, but no evidence has been adduced in favour of a specific recognition sequence in those mRNAs which are used as primers by the virus. It is perhaps more likely that some physical constraints on the transcriptase complex within the cell nucleus may impede the proper regulation of transcription in certain host cells.

VI. Amantadine-Resistant Mutants

A number of drugs will inhibit the replication of influenza viruses in cell culture, and several potential chemotherapeutic agents have been identified (Oxford, 1975; Oxford and Williams, 1976). However, genetic studies on drug resistance have been almost totally confined to amantadine, (1-adamantanamine) and its derivatives

(Field, 1983), first studied by Davies *et al.* (1964). This drug inhibits influenza virus replication at a very early stage, before primary transcription (Skehel *et al.*, 1977), but although "uncoating" seems to be inhibited (Kato and Eggers, 1969; Koff and Knight, 1979), the exact step which is affected has not been defined. Amantadine is concentrated intracellularly in lysosomes, where it has been found to raise the pH (Ohkuma and Poole, 1978), but the antiviral properties of the drug appear to depend upon its presence in the extracellular medium rather than within the cell (Richman *et al.*, 1981).

The development of drug-resistant mutants was observed in cell cultures (Cochran *et al.*, 1965) and in amantadine-treated mice (Oxford *et al.*, 1970 a, b), but resistance to amantadine or the related compound cyclooctylamine was found to vary greatly between influenza virus strains; in a plaque reduction test, the 50% value was 0.1 μg/ml for the A/Singapore/1/57 strain compared to 20 μg/ml for A/PR/8/34, and resistance varied in genetic crosses independently of the surface antigens (Tuckova *et al.*, 1973; Appleyard and Maber, 1975; Appleyard, 1977).

The RNA segment carrying the gene for amantadine sensitivity was located by a comparison of the genetic composition of recombinant strains with their response to amantadine. RNA segment 7 alone appeared to confer resistance or sensitivity to the drug (Hay *et al.*, 1979; Lubeck *et al.*, 1978), although as this segment generates three mRNAs in infected cells it is not clear which polypeptide function is involved. However, Zvonarjev and Ghendon (1980) studied the influence of rimantadine on an RNA transcriptase reaction catalyzed by influenza virus RNP prepared from deoxycholate-disrupted virions. Addition of M protein to this reaction depressed *in vitro* transcription by 40%, and addition of 200 μg rimantadine further depressed the activity by a further 40%. This effect of rimantadine was not observed if the M protein added *in vitro* was derived from a rimantadine-resistant mutant, suggesting that the drug may interact in some way with the matrix protein (Zvonarjev and Ghendon, 1980).

Recently Bukrinskaya *et al.* (1982 a) identified two intermediates in virus uncoating; subviral particles (containing RNP's and M protein) were found in the nucleus-associated cytoplasm and RNP's (free of M protein) in the nucleoplasm. Rimantadine blocked the appearance of RNP's in both nucleus and cytoplasm of cells infected with rimantadine-sensitive viruses, but in cells infected with rimantadine-resistant strains RNP's were uncoated normally. It was concluded that rimantadine blocks the second stage of influenza virus uncoating, removal of M protein from the RNP's (Bukrinskaya *et al.*, 1982 b).

In contrast to these results regarding the role of the M protein, Scholtissek and Faulkner (1979) have suggested that amantadine resistance may not be confined to a single gene product or mechanism. Differences were found in a single virus strain, *e.g.* fowl plague virus was amantadine-sensitive as measured by haemagglutinin yield in a single cycle, but resistant as measured in a plaque reduction test. A series of strains and recombinants prepared using *ts* G mutants of fowl plague virus was analyzed. The haemagglutinin gene was assocated with transfer of resistance in the haemagglutinin yield test, but not in the plaque reduction test, where segments 5 and 6 transferred sensitivity, and segment 7 occasionally transferred resistance. It was concluded that the correct gene constellation, rather than transfer of a single

gene, was required for amantadine resistance as was found with various pathogenicity and host range markers.

VII. General Conclusions

The recognition in the mid-seventies that the influenza virus genome consists of eight distinct RNA segments provided considerable stimulus to genetic studies of the virus. Compared with the investigation of genetic systems in other negative-strand viruses, such as vesicular stomatitis virus (Pringle, 1982), the influenza virus geneticist has few problems in identification of the locus of a mutation, and can where necessary determine the precise nucleotide sequence alteration (Robertson *et al.*, 1983). Indeed, the path is clear towards the creation of site-specific mutations in particular influenza virus genes cloned in M13 (Winter *et al.*, 1982) although there are problems still to be overcome in reintroducing such genes into the virus-host cell system.

Despite these obvious advantages, the segmented genome has attendant complications in the ease with which gene reassortment can result in extragenic suppression of a mutant phenotype. This type of genetic interaction is hard to recognize, and may to some extent underlie the difficulty in establishing specific phenotypes for influenza virus mutants with lesions in particular genome segments. However another obvious feature of phenotypic characterization is pleiotropism, reflecting the close interaction of certain virus-specific proteins with each other. Pleiotropic effects have been found between groups of proteins such as PB1, PB2 and NS_1, NP and M, and HA and NA, and were clearly recognized with influenza virus mutants even before the assignment of *ts* lesions to individual RNA segments had been made (Hightower and Bratt, 1977). The largest number of available *ts* mutants of influenza virus have defects in RNA segment 1, but within this group the range of phenotypes is considerable.

Two approaches should prove valuable in further studies in which mutants of influenza virus are employed. First, the primary lesion in a number of non-leaky mutants with lesions in a particular influenza virus genome RNA segment should be established by careful investigation of each of the established stages in the replication cycle. The use of virus cores to compare *in vitro* virion transcriptase activities of *ts* mutant and wild-type viruses should greatly assist interpretation of results from incubations at the restrictive temperature (Ulmanen *et al.*, 1983). Methods recently developed which employ cloned copies of individual RNA segments (Thierry and Danos, 1982) can be adopted to allow more direct analysis of the RNA phenotype than has hitherto been possible. Finally, the application of two-dimensional gel electrophoretic techniques combined with isoelectric focusing or non-equilibrium pH gradient electrophoresis offers considerable advantages in the characterization of polypeptide alterations in mutant virus-infected cells. This technique has greatly clarified the characterization of the P polypeptides, which are difficult to resolve in single-dimension electrophoresis (Horisberger, 1980), and can, for example, easily detect the single charge change on the NS_2 polypeptide specified by *ts* C47 (C. R. Penn, unpublished observations).

The second approach of value in mutant characterization is to study in considerable detail selected single mutants which have a phenotype of particular interest. Mutants which fail to haemagglutinate or which are known to be defective in 5′-terminal cap recognition, for example, could well repay a thorough investigation of their protein structural alterations.

Finally, in reviewing this field it is surprizing to find that more effort has not been made to obtain mutants with lesions in segments 7 and 8, where few examples are available for study. It is likely that an understanding of the functions of up to four non-structural proteins specified by these two segments will only come through a study of their conditional lethal mutants.

Acknowledgements

The preparation of this review was greatly assisted by the following individuals who provided information or manuscripts prior to publication: J. W. Almond, C.-M. Chu, Y. Z. Ghendon, A. J. Hay, A. P. Kendal, R. M. Krug, R. A. Lamb, H. F. Maassab, B. R. Murphy, R. Rott, C. Scholtissek, A. Sugiura, and G. P. Winter. I am grateful to my colleagues, Steve Inglis and Jamie Robertson, for many useful discussions, and especially to Charles Penn, John McCauley, Mike Carter, and Donald Smith, who were good enough to read and criticize the manuscript. Finally, I wish to thank Mary Wright for so skilfully and tirelessly rendering my manuscript into typescript. Our work in Cambridge is supported by a programme grant from the Medical Research Council.

References

Alexandrova, G. I., Garmashova, L. M., Golubev, D. B., Koljak, C. I., Medvedeva, T. E., Polezhaev, F. I.: The experience of selection of thermosensitive recombinants of influenza virus type 2. Vopr. Virusol. *4,* 342–346 (1979).

Alexandrova, G. I., Smorodintsev, A. A.: Obtaining of an additionally attenuated vaccinating cryophilic influenza strain. Rev. roum. Intramicrobiol. *2,* 179–186 (1965).

Allen, H., McCauley, J., Waterfield, M., Gething, M.-J.: Influenza virus segment 7 has the coding capacity for two polypeptides. Virology *107,* 548–551 (1980).

Almond, J. W.: Temperature-sensitive mutants of fowl plague virus. Ph. D. Thesis, University of Cambridge (1978).

Almond, J. W.: A single gene determines the host range of influenza. Nature *270,* 617–618 (1977).

Almond, J. W., Barry, R. D.: A single gene controlling the host range of fowl plague virus. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 675–684. New York: Academic Press 1978.

Almond, J. W., Barry, R. D.: Genetic recombination between two strains of fowl plague virus: construction of genetic maps. Virology *92,* 407–415 (1979).

Almond, J. W., McGeoch, D., Barry, R. D.: Method for assigning temperature-sensitive mutations of influenza viruses to individual segments of the genome. Virology *81,* 62–73 (1977).

Almond, J. W., McGeoch, D., Barry, R. D.: Temperature-sensitive mutants of fowl plague virus: isolation and genetic characterization. Virology *92,* 416–427 (1979).

Ames, B. N., Whitfield, H. J., jr.: Frame-shift mutagenesis in Salmonella. Cold Spring Harbor Symp. Quant. Biol. *31,* 221–225 (1966).

Anisimova, E., Ghendon, Y., Markushin, S.: Ultrastructural changes in cells induced by temperature-sensitive mutants of fowl plague virus at permissive and non-permissive temperature. J. Gen. Virol. *47,* 11–18 (1980).

Anisimova, E., Tuckova, E., Vonka, V., Zavadova, A.: Ultrastructural changes induced by influenza viruses in permissive and non-permissive cells. Virology *77,* 330–336 (1977).

Appleyard, G.: Amantadine-resistance as a genetic marker for influenza viruses. J. Gen. Virol. *36,* 249–255 (1977).

Appleyard, G., Maber, H. B.: Plaque by influenza viruses in presence of trypsin. J. Gen. Virol. *25,* 351–357 (1974).

Appleyard, G., Maber, H. B.: A plaque assay for the study of influenza virus inhibitors. J. Antimicrob. Chemother. *1* (Suppl.), 49–53 (1975).

Barrett, T., Wolstenholme, A. J., Mahy, B. W. J.: Transcription and replication of influenza virus RNA. Virology *98,* 211–225 (1979).

Barry, R. D.: The multiplication of influenza virus. II. Multiplicity reactivation of ultraviolet irradiated virus. Virology *14,* 398–405 (1961).

Barry, R. D., Mahy, B. W. J.: The influenza virus genome and its replication. Brit. Med. Bull. *35,* 39–46 (1979).

Bean, W. J., Simpson, R. W.: Primary transcription of the influenza virus genome in permissive cells. Virology *56,* 646–651 (1973).

Bean, W. J., Simpson, R. W.: Associated transcriptase activity of influenza recombinant and mutant strains. J. Virol. *16,* 516–525 (1975).

Bean, W. J., Simpson, R. W.: Transcriptase activity and genome composition of defective influenza virus. J. Virol. *18,* 365–369 (1976).

Bean, W. J., jr., Webster, R. G.: Phenotypic properties associated with influenza genome segments. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 685–692. New York: Academic Press 1978.

Bentley, D. R., Brownlee, G. G.: Sequence of the N2 neuraminidase from influenza virus A/NT/60/68. Nucleic Acids Res. *10,* 5033–5042 (1982).

Bishop, D. H. L., Huddleston, J. A., Brownlee, G. G.: The complete sequence of RNA segment 2 of influenza A/NT/60/68 and its encoded P_1 protein. Nucleic Acids Res. *10,* 1335–1343 (1982 a).

Bishop, D. H. L., Jones, K. L., Huddleston, J. A., Brownlee, G. G.: Influenza A virus evolution: complete sequences of influenza A/NT/60/68 RNA segment 3 and its predicted acidic P polypeptide compared with those of influenza A/PR/8/34. Virology *120,* 481–489 (1982 b).

Blaas, D., Patzelt, E., Kuechler, E.: Cap-recognizing protein of influenza virus. Virology *116,* 339–348 (1982 a).

Blaas, D., Patzelt, E., Kuechler, E.: Identification of the cap binding protein of influenza virus. Nucleic Acids Res. *10,* 4803–4812 (1982 b).

Brand, C., Palese, P.: Sequential passage of influenza virus in embryonated egg or tissue culture: emergence of mutants. Virology *107,* 424–433 (1980).

Briedis, D. J., Conti, G., Munn, E. A., Mahy, B. W. J.: Migration of influenza virus-specific polypeptides from cytoplasm to nucleus of infected cells. Virology *111,* 154–164 (1981).

Bukrinskaya, A. G., Vorkunova, N. K., Kornilayeva, G. V., Narmanbetova, R. A., Vorkunova, G. K.: Influenza virus uncoating in infected cells and effect of rimantadine. J. Gen. Virol. *60,* 49–59 (1982 a).

Bukrinskaya, A. G., Vorkunova, N. K., Pushkarskaya, N. L.: Uncoating of a rimantadine-resistant variant of influenza virus in the presence of rimantadine. J. Gen. Virol. *60,* 61–66 (1982 b).

Burge, B. W., Pfefferkorn, E. R.: Conditional lethal mutants of an animal RNA virus. Virology *24,* 126–128 (1964).

Burnet, F. M., Lind, P. E.: Recombination of characters between two influenza strains. Austral. J. Sci. *12,* 109–110 (1949).

Burnet, F. M., Lind, P. E.: Studies on recombination with influenza viruses in the chick embryo. III. Reciprocal genetic interaction between two influenza virus strains. Austral. J. Exp. Biol. *30,* 469–477 (1952).

Caliguiri, L. A., Compans, R. W.: Analysis of the *in vitro* product of an RNA-dependent RNA polymerase isolated from influenza virus-infected cells. J. Virol. *14,* 191–197 (1974).

Caliguiri, L. A., Holmes, K. V.: Host-dependent restriction of influenza virus maturation. Virology *92,* 15–30 (1979).

Carter, M. J., Mahy, B. W. J.: Incomplete avian influenza virus contains a defective non-interfering component. Arch. Virol. *71,* 13–25 (1982 a).

Carter, M. J., Mahy, B. W. J.: Synthesis of RNA segments 1–3 during generation of incomplete influenza A (fowl plague) virus. Arch. Virol. *73,* 109–119 (1982 b).

Caton, A. J., Robertson, J. S.: Structure of the host-derived sequences present at the 5′ ends of influenza virus mRNA. Nucleic Acids Res. *8,* 2591–2603 (1980).

Chu, C.-M., Tian, S.-F., Ren, G.-F., Zhang, Y.-M., Zhang, L.-X., Liu, G.-Q.: Occurrence of temperature-sensitive influenza A viruses in nature. J. Virol. *41,* 353–359 (1982).

Cochran, K. W., Maassab, H. E., Tsunoda, A., Berlin, B. S.: Studies on the antiviral activity of amantadine hydrochloride. Ann. N.Y. Acad. Sci. *130,* 432–440 (1965).

Compans, R. W., Caliguiri, L. A.: Isolation and properties of an RNA polymerase from influenza virus-infected cells. J. Virol. *11,* 441–448 (1973).

Cox, N. J., Kendal, A. P., Maassab, H. F., Scholtissek, C., Spring, S. B.: Genetic synergism between matrix protein and polymerase protein required for temperature-sensitivity of the cold-adapted influenza A/Ann Arbor/6/60 mutant virus. In: The Replication of Negative Strand Viruses (Bishop, D. H., Compans, R. W., eds.), 405–413. Elsevier/North-Holland 1981 a.

Cox, N. J., Konnecke, I., Kendal, A. P., Maassab, H. F.: Genetic and biochemical analysis of the A/Ann Arbor/6/60 cold-adapted mutant. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., ed.), 639–652. New York: Academic Press 1981 b.

Cremer, K. J., Bodemer, M., Summers, W. P., Summers, W. C., Gesteland, R. F.: *In vitro* suppression of UAG and UGA mutants in the thymidine kinase gene of herpes simplex virus. Proc. Natl. Acad. Sci. U.S.A. *76,* 430–434 (1979).

Dales, S., Pons, M. W.: Penetration of influenza examined by means of virus aggregates. Virology *69,* 278–286 (1976).

Davies, P., Barry, R. D.: Nucleic acid of influenza virus. Nature *211,* 384–387 (1966).

Davies. W. L., Grunert, R. R., Haff, R. F., McGahen, J. W., Neumayer, E. M., Paulshock, M., Watts, J. C., Wood, T. R., Hermann, E. C., Hoffmann, C. E.: Antiviral activity of 1-adamantanamine (amantadine). Science *144,* 862–863 (1964).

Desselberger, U., Palese, P.: Molecular weights of RNA segments of influenza A and B viruses. Virology *88,* 394–399 (1978).

Desselberger, U., Racaniello, V. R., Zazra, J. J., Palese, P.: The 3′ end 5′-terminal sequences of influenza A, B, and C virus RNA segments are highly conserved and show partial inverted complementarity. Gene *8,* 315–328 (1980).

Dhar, R., Chanock, R. M., Lai, C.-J.: Nonviral oligonucleotides at the 5′-terminus of cytoplasmic influenza viral mRNA deduced from cloned complete genomic sequences. Cell *21,* 495–500 (1980).

Dimmock, N. J.: New virus-specific antigens in cells infected with influenza virus. Virology *39,* 224–234 (1969).

Donald, H. B., Isaacs, A.: Counts of influenza virus particles. J. Gen. Microbiol. *10,* 457–464 (1954).

Duesberg, P. H., Robinson, W. S.: On the structure and replication of influenza virus. J. Mol. Biol. *25,* 383–405 (1967).

Fenner, F., Sambrook, J. F.: The genetics of animal viruses. Ann. Rev. Microbiol. *18,* 47–94 (1964).

Field, H. J.: The problem of drug-induced resistance in viruses. In: The 5th Beecham Colloquium "Problems of Antiviral Chemotherapy" (Stuart-Harris, Ch., Oxford, J. S., eds.), 71–107. London: Academic Press 1983.

Fields, B. N.: Genetics of reovirus. Curr. Topics Microbiol. Immunol. *91,* 1–24 (1981).

Fields, B. N., Joklik, W. K.: Isolation and preliminary genetic and biochemical characterization of temperature-sensitive mutants of reovirus. Virology *37,* 335–342 (1969).

Fields, S., Winter, G.: Nucleotide-sequence heterogeneity and sequence rearrangements in influenza virus cDNA. Gene *15,* 207–214 (1981 a).

Fields, S., Winter, G.: Influenza virus A/PR/8/34 genes: sequencing by a shotgun approach. In: Genetic Variation Among Influenza Viruses (Nayak, D., Fox, C. F., eds.). ICN-UCLA Symp. Mol. Cell. Biol. *21,* 55–64. New York: Academic Press 1981 b.

Fields, S., Winter, G.: Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structure of a small viral RNA segment. Cell *28,* 303–313 (1982).
Fields, S., Winter, G., Brownlee, G. G.: Structure of the neuraminidase gene in human influenza virus A/PR/8/34. Nature *290,* 213–217 (1981).
Fowler, R. G., Degnen, G. E., Cox, E. C.: Mutational specificity of a conditional Escherichia coli mutator, mut D5. Molec. Gen. Genet. *133,* 179–191 (1974).
Freese, E.: The difference between spontaneous and base analogue induced mutations of phage T4. Proc. Nat. Acad. Sci. U.S.A. *45,* 622–633 (1959).
Freese, E., Bantz-Freese, E., Bantz, E.: Hydroxylamine as a mutagenic and inactivating agent. J. Mol. Biol. *3,* 133–143 (1961).
Ghendon, Y. Z.: Conditional-lethal mutants of animal viruses. Prog. Med. Virol. *14,* 68–122 (1972).
Ghendon, Y., Klimov, A., Blagoveshenskaya, O., Genkina, D.: Investigation of recombinants of human influenza and fowl plague viruses. J. Gen. Virol. *43,* 183–191 (1979 a).
Ghendon, Y. Z., Markushin, S. G.: Studies on mutation lesions and physiology of fowl plague virus *ts* mutants. Phil. Trans. R. Soc. (Lond.) *B 288,* 383–392 (1980).
Ghendon, Y. Z., Markushin, S. G., Blagovezhenskaya, O. V., Genkina, D. B.: Study of fowl plague virus RNA synthesis in temperature sensitive mutants. Virology *66,* 454–463 (1975).
Ghendon, Y. Z., Markushin, S. G., Ginzburg, V. P., Hay, A.: Functional defects of a fowl plague virus temperature-sensitive mutant having mutation in the neuraminidase. Arch. Virol. *75,* 55–70 (1983 a).
Ghendon, Y. Z., Markushin, S. G., Klimov, A. I., Hay, A. J.: Studies of fowl plague virus temperature-sensitive mutants with defects in transcription. J. Gen. Virol. *63,* 103–111 (1982 a).
Ghendon, Y. Z., Markushin, S. G., Klimov, A. I., Lotte, V. D., Ginzburg, V. P.: Studies of a temperature-sensitive fowl plague virus *ts* mutant having a mutation in gene 7 coding for the M protein. J. Gen. Virol. *64,* 291–304 (1983 b).
Ghendon, Y., Markushin, S., Lisovskaya, K., Penn, C. R., Mahy, B. W. J.: Extragenic suppression of *ts* phenotype during recombination between *ts* mutants of two fowl plague virus strains with a *ts* mutation in gene 1. J. Gen. Virol. *62,* 239–248 (1982 b).
Ghendon, Y. Z., Markushin, S. G., Marchenko, A. T., Sitnikov, B. S., Ginsburg, V. P.: Biochemical characteristics of fowl plague virus *ts* mutants. Virology *55,* 305–319 (1973).
Ghendon, Y., Tuckova, E., Vonka, V., Klimov, A., Ginzburg, V., Markushin, S.: Replication of two influenza virus strains and a recombinant in HEF and LEP cells. J. Gen. Virol. *44,* 179–186 (1979 b).
Ghenkina, D. B., Ghendon, Y. Z.: Recombination and complementation between orthomyxoviruses under conditions of abortive infection. Acta Virol. *23,* 97–106 (1979).
Glass, S. E., McGeoch, D., Barry, R. D.: Characterization of the mRNA of influenza virus. J. Virol. *16,* 1435–1443 (1975).
Glass, R. E., Nene, V., Hunter, M. G.: Informational suppression as a tool for the investigation of gene structure and function. Biochem. J. *203,* 1–13 (1982).
Gopinathan, K. P., Weymouth, L. A., Kunkel, T. A., Loeb, L. A.: Mutagenesis *in vitro* by DNA polymerase from an RNA virus. Nature *278,* 857–859 (1979).
Gordon, M. P., Staehelin, M.: Studies on the incorporation of 5-fluorouracil into a virus nucleic acid. Biochim. Biophys. Acta *36,* 351–361 (1959).
Gregoriades, A., Hirst, G. K.: Mechanism of influenza recombination. III. Biochemical studies of temperature-sensitive mutants belonging to different recombination groups. Virology *69,* 81–92 (1976).
Harms, E., Rohde, W., Bosch, F., Scholtissek, C.: Biochemical studies on influenza viruses. II. Assignment of gene functions to RNA segments 5, 7, and 8 of fowl plague virus and virus N. Virology *86,* 413–422 (1978).
Hartman, P. E., Roth, J. R.: Mechanisms of suppression. Advances in Genetics *17,* 1–105 (1973).
Hastie, N. D., Mahy, B. W. J.: RNA-dependent RNA polymerase in nuclei of cells infected with influenza virus. J. Virol. *12,* 951–961 (1973).
Hay, A. J., Bellamy, A. R., Abraham, G., Skehel, J. J., Brand, C. M., Webster, R. G.: Procedures for characterization of the genetic material of candidate vaccine strains. Developments in Biological Standardization *39,* 15–24 (1977 a).

Hay, A.J., Kennedy, N.I.T., Skehel, J.J., Appleyard, G.: The matrix protein gene determines amantadine-sensitivity of influenza viruses. J. Gen. Virol. *42,* 189–191 (1979).

Hay, A.J., Lomniczi, B., Bellamy, A.R., Skehel, J.J.: Transcription of the influenza virus genome. Virology *83,* 337–355 (1977 b).

Hay, A.J., Skehel, J.J., McCauley, J.: Structure and synthesis of influenza virus complementary RNAs. Phil. Trans. Royal Soc. (Lond.) *B 288,* 341–348 (1980).

Hay, A.J., Skehel, J.J., McCauley, J.: Characterization of influenza virus RNA complete transcripts. Virology *116,* 517–522 (1982).

Heller, E., Scholtissek, C.: Evidence for intracistronic complementation of the product of the influenza virus gene *Ptra* (P3 of fowl plague virus). J. Gen. Virol. *49,* 133–139 (1980).

Henle, W., Liu, O.C.: Studies on host-virus interactions in the chick embryo-influenza virus system. VI. Evidence for multiplicity reactivation of inactivated virus. J. Exp. Med. *94,* 305–322 (1951).

Hightower, L.E., Bratt, M.A.: Genetics of orthomyxoviruses. In: Comprehensive Virology (Fraenkel Conrat, H., Wagner, R.R., eds.), Vol. 9, 535–598. New York: Plenum Press 1977.

Hirst, G.K.: Genetic recombination with Newcastle disease virus, polioviruses and influenza. Cold Spring Harbor Symp. Quant. Biol. *27,* 303–308 (1962).

Hirst, G.K.: Mechanism of influenza virus recombination. I. Factors influencing recombination rates between temperature-sensitive mutants of strain WSN and the classification of mutants into complementation-recombination groups. Virology *55,* 81–93 (1973).

Hirst, G.K., Gotlieb, T.: The experimental combination of forms of virus. V. Alterations in the virulence of neurotropic influenza virus as a result of mixed infection. Virology *1,* 221–235 (1955).

Hirst, G.K., Pons, M.W.: Mechanism of influenza recombination. II. Virus aggregation and its effect on plaque formation by so-called non-infectious virus. Virology *56,* 620–631 (1973).

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., Vande Pol, S.: Rapid evolution of RNA genomes. Science *215,* 1577–1585 (1982).

Horisberger, M.A.: The large P proteins of influenza A viruses are composed of one acidic and two basic polypeptides. Virology *107,* 302–305 (1980).

Horisberger, M.A.: Identification of a catalytic activity of the large basic P polypeptide of influenza virus. Virology *120,* 279–286 (1982).

Horisberger, M.A., De Staritzky, C.: Two dimensional gel analysis of the influenza B proteins of the transcriptase complex. FEMS Microbiol. Lett. *12,* 95–98 (1981).

Hoyle, L.: The influenza viruses. Virology Monographs, Vol. 4. Wien-New York: Springer 1968.

Huddleston, J.A., Brownlee, G.G.: The sequence of the nucleoprotein gene of human influenza A virus, strain A/NT/60/68. Nucleic Acids Res. *10,* 1029–1038 (1982).

Inglis, S.C.: Characterization of subgenomic polyadenylated mRNAs encoded by genome RNA segment 7 of influenza virus. In: The Replication of Negative Strand Viruses (Bishop, D.H.L., Compans, R.W., eds.), 261–267. New York: Elsevier/North-Holland 1981.

Inglis, S.C., Barrett, T., Brown, C.M., Almond, J.W.: The smallest genome RNA segment of influenza virus contains two genes that may overlap. Proc. Natl. Acad. Sci. U.S.A. *76,* 3790–3794 (1979).

Inglis, S.C., Brown, C.M.: Spliced and unspliced RNAs encoded by virion RNA segment 7 of influenza virus. Nucleic Acids Res. *9,* 2727–2740 (1981).

Inglis, S.C., Carroll, A.R., Lamb, R.A., Mahy, B.W.J.: Polypeptides specified by the influenza virus genome. 1. Evidence for eight distinct gene products specified by fowl plague virus. Virology *74,* 489–503 (1976).

Inglis, S.C., Conti, G., Mahy, B.W.J.: Control of influenza virus polypeptide synthesis. In: Negative Strand Viruses and the Host Cell (Mahy, B.W.J., Barry, R.D., eds.), 239–248. London: Academic Press 1978.

Inglis, S.C., Gething, M.-J., Brown, C.M.: Relationship between the messenger RNAs transcribed from two overlapping genes of influenza virus. Nucleic Acids Res. *8,* 3575–3589 (1980).

Inglis, S.C., Mahy, B.W.J.: Polypeptides specified by the influenza virus genome. 3. Control of synthesis in infected cells. Virology *95,* 154–164 (1979).

Inglis, S.C., McGeoch, D.J., Mahy, B.W.J.: Polypeptides specified by the influenza virus genome. 2. Assignment of protein coding functions to individual genome segments by *in vitro* translation. Virology *78,* 522–536 (1977).

Isaacs, A., Donald, H. B.: Particle counts of haemagglutinating viruses. J. Gen. Microbiol. *12,* 241–247 (1955).

Israel, A.: Temperature-dependent host range mutant of fowl plague virus (FPV). Virology *105,* 1–12 (1980 a).

Israel, A.: Productive and abortive infection of L cells by fowl plague virus (FPV): comparison of *in vivo* and *in vitro* translation products of the virus mRNAs. J. Gen. Virol. *47,* 473–483 (1980 b).

Israel, A., Semmel, M., Huppert, J.: Host range mutant of fowl plague virus (FPV): comparison of the genome and virus proteins. Virology *68,* 503–509 (1975).

Kato, N., Eggers, H. J.: Inhibition of uncoating of fowl plague virus by 1-adamantanamine hydrochloride. Virology *37,* 632–641 (1969).

Kawakami, K., Ishihama, H.: RNA polymerase of influenza virus III. Isolation of RNA-polymerase-RNA complexes from influenza PR8. J. Biochem. (Tokyo) (in press, 1983).

Kendal, A. P., Cox, N. J., Galphin, J. C., Maassab, H. F.: Comparative studies of wild-type and cold-mutant (temperature sensitive) influenza viruses: independent segretation of temperature-sensitivity of virus replication from temperature-sensitivity of virion transcriptase activity during recombination of mutant A/Ann Arbor/6/60 with wild-type H3N2 strains. J. Gen. Virol. *44,* 443–456 (1979).

Kendal, A. P., Galphin, J. C., Palmer, E.: Replication of influenza virus at elevated temperatures: production of virus-like particles with reduced matrix protein content. Virology *76,* 186–195 (1977).

Kendal, A. P., Maassab, H. F., Alexandrova, G. I., Ghendon, Y.: Development of cold-adapted recombinant live, attenuated influenza A vaccines in the U.S.A. and U.S.S.R. Antiviral Res. *1,* 339–365 (1981).

Kilbourne, E. D.: Influenza virus genetics. Progr. Med. Virol. *5,* 79–126 (1963).

King, A. M. Q., McCahon, D., Slade, W. R., Newman, J. W. I.: Recombination in RNA. Cell *29,* 921–928 (1982).

Klenk, H.-D., Rott, R., Orlich, M., Blödorn, J.: Activation of influenza A viruses by trypsin treatment. Virology *68,* 426–439 (1975).

Koennecke, I., Boschek, C. B., Scholtissek, C.: Isolation and properties of a temperature-sensitive mutant (*ts* 412) of an influenza A virus recombinant with a *ts* lesion in the gene coding for the nonstructural protein. Virology *110,* 16–25 (1981).

Koff, W. C., Knight, V.: Inhibition of influenza virus incoating by rimantadine hydrochloride. J. Virol. *31,* 261–263 (1979).

Krug, R. M.: Priming of influenza viral RNA transcription by capped heterologous RNAs. Curr. Topics Microbiol. Immunol. *93,* 125–149 (1981).

Krug, R. M., Broni, B. A., Bouloy, M.: Are the 5′ ends of influenza viral mRNAs synthesized *in vitro* donated by host mRNAs? Cell *18,* 329–334 (1979).

Krug, R. M., Ueda, M., Palese, P.: Temperature-sensitive mutants of influenza WSN virus defective in virus-specific RNA synthesis. J. Virol. *16,* 790–796 (1975).

Lamb, R. A., Choppin, P. W.: Segment 8 of the influenza virus genome is unique in coding for two polypeptides. Proc. Natl. Acad. Sci. U.S.A. *76,* 4908–4912 (1979).

Lamb, R. A., Choppin, P. W.: Identification of a second protein (M_2) encoded by RNA segment 7 of influenza virus. Virology *112,* 729–737 (1981).

Lamb, R. A., Lai, C. J.: Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. Cell *21,* 475–485 (1980).

Lamb, R. A., Lai, C. J.: Conversation of the influenza membrane protein (M1) amino acid sequence and an open reading frame of RNA segment 7 encoding a second protein (M2) in H1N1 and H3N2 strains. Virology *112,* 746–751 (1981 a).

Lamb, R. A., Lai, C. J.: Interrupted mRNA(s) and overlapping genes in influenza virus. In: The Replication of Negative Strand Viruses (Bishop, D. H. L., Compans, R. W., eds.), 251–259. New York: Elsevier/North-Holland 1981 b.

Lamb, R. A., Lai, C. J., Choppin, P. W.: Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping protein. Proc. N.A.S. (U.S.) *78,* 4170–4174 (1981).

Lohmeyer, J., Klenk, H.-D.: A mutant of influenza virus with a temperature-sensitive defect in the post-translational processing of the haemagglutinin. Virology *93*, 134–145 (1979).

Loveless, A.: Possible relevance of 0–6 alkylation of deoxyguanosine to the mutagenicity and carcinogenicity of nitrosamines and nitrosamides. Nature *223*, 206–207 (1969).

Lubeck, M. D., Schulman, J. L., Palese, P.: Susceptibility of influenza A viruses to amantadine is influenced by the gene coding for M protein. J. Gen. Virol. *28*, 710–716 (1978).

Maassab, H. F.: Adaptation and growth characteristics of influenza virus at 25 °C. Nature *219*, 645–646 (1967).

Maassab, H. F., Monto, A. S., DeBorde, D. C., Cox, N. J., Kendal, A. P.: Development of cold recombinants of influenza virus as live virus vaccines. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., ed.), 617–637. ICN-UCLA Symposium. New York: Academic Press 1981.

Maassab, H. F., Spring, S. B., Kendal, A. P., Monto, A. S.: Biologic characteristic of influenza virus recombinants derived at sub-optimal temperatures. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 721–732. London: Academic Press 1978.

Mackenzie, J. S.: Virulence of temperature-sensitive mutants of influenza virus. Br. Med. J. *3*, 757–758 (1969).

Mackenzie, J. S.: Isolation of temperature-sensitive mutants and the construction of a preliminary genetic map for influenza virus. J. Gen. Virol. *6*, 63–75 (1970 a).

Mackenzie, J. S.: Studies with temperature-sensitive mutants of influenza virus. In: The Biology of Large RNA Viruses (Barry, R. D., Mahy, B. W. J., eds.), 507–534. London: Academic Press 1970 b.

Mackenzie, J. S., Dimmock, N. J.: A preliminary study of physiological characteristics of temperature-sensitive mutants of influenza virus. J. Gen. Virol. *19*, 51–63 (1973).

Mahy, B. W. J., Barrett, T., Briedis, D. J., Brownson, J. M., Wolstenholme, A. J.: Influence of the host cell on influenza virus replication. Phil. Trans. R. Soc. (Lond.) *B 288*, 349–357 (1980).

Mahy, B. W. J., Barrett, T., Nichol, S. T., Penn, C. R., Wolstenholme, A. J.: Analysis of the functions of influenza virus genome RNA segments by use of temperature-sensitive mutants of fowl plague virus. In: The Replication of Negative Strand Viruses (Bishop, D. H. L., Compans, R. W., eds.), 379–387. Elsevier/North-Holland 1981 a.

Mahy, B. W. J., Bromley, P. A.: *In vitro* product of a ribonucleic acid polymerase induced by influenza virus. J. Virol. *6*, 259–268 (1970).

Mahy, B. W. J., Hastie, N. D., Raper, R. H., Brownson, J. M. T., Carroll, A. R.: RNA polymerase activities of nuclei from influenza virus-infected cells. In: Negative Strand Viruses (Mahy, B. W. J., Barry, R. D., eds.), 445–467. London: Academic Press 1975.

Mahy, B. W. J., Penn, C. R., Nichol, S. T., Briedis, D. J., Barrett, T.: Influenza virus-specific proteins regulating genome transcription. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., ed.), 113–125. (ICN-UCLA Symposia on Molecular and Cellular Biology, Vol. 21.) Academic Press 1981 b.

Mark, G. E., Taylor, J. M., Broni, R., Krug, R. M.: Nuclear accumulation of influenza viral RNA transcripts and the effects of cycloheximide, actinomycin D, and α-amanitin. J. Virol. *29*, 744–752 (1979).

Markushin, S. G., Ghendon, Y. Z.: Genetic classification and biological properties of temperature-sensitive mutants of fowl plague virus. Acta Virol. *17*, 369–376 (1973).

Markushin, S. G., Isachenko, V. A., Molibug, E. V., Zakstelskaya, L. Y., Ghendon, Y. Z.: Gene homology and antigenic specificity of envelope proteins of avian influenza virus strains possessing haemagglutinin Hav1. Acta Virol. *25*, 65–70 (1981).

Massicot, J. G., Murphy, B. R., Thierry, F., Markoff, L., Huang, K.-Y., Chanock, R. M.: 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 *ts* lesions with a predictable level of attenuation. Virology *101*, 242–249 (1980 a).

Massicot, J. G., Murphy, B. R., van Wyke, K., Huang, K.-Y., Chanock, R. M.: *ts* P1 and P3 genes are responsible for satisfactory level of attenuation of *ts*-1A2 recombinants bearing H1N1 of H3N2 surface antigens of influenza A virus. Virology *106*, 187–190 (1980 b).

Massicot, J. G., van Wyke, K., Chanock, R. M., Murphy, B. R.: Evidence for intrasegmental complementation between two influenza A viruses having *ts* mutations on their P1 genes. Virology *117*, 496–500 (1982).

McCauley, J. W., Mahy, B. W. J., Inglis, S. C.: Nucleotide sequence of fowl plague virus RNA segment 7. J. Gen. Virol. *58,* 211–215 (1982).

McGeoch, D. J.: Structural analysis of animal virus genomes. J. Gen. Virol. *55,* 1–16 (1981).

McGeoch, D., Fellner, P., Newton, C.: The influenza virus genome consists of eight distinct RNA species. Proc. Natl. Acad. Sci. U.S.A. *73,* 3045–3049 (1976).

McGeoch, D., Kitron, N.: Influenza virion RNA-dependent RNA polymerase: stimulation by guanosine and related compounds. J. Virol. *15,* 686–695 (1975).

Mills, J., Chanock, R. M.: Temperature-sensitive mutants of influenza virus. I. Behaviour in tissue culture and in experimantal animals. J. Infect. Dis. *123,* 145–157 (1971).

Morrongiello, M. P., Dales, S.: Characterization of cytoplasmic inclusions formed during influenza/WSN virus infection of chick embryo fibroblasts. Intervirology *8,* 281–293 (1977).

Mowshowitz, S. L.: P1 is required for initiation of cRNA synthesis in WSN influenza virus. Virology *91,* 493–495 (1978).

Mowshowitz, S. L.: RNA synthesis of temperature-sensitive mutants of WSN influenza virus. In: Replication of Negative Strand Viruses (Bishop, D. W. L., Compans, R. W., eds.), 317–323. Elsevier/North-Holland 1981.

Mowshowitz, S. L., Ueda, M.: Temperature-sensitive virion transcriptase activity in mutants of WSN influenza virus. Arch. Virol. *52,* 135–141 (1976).

Murphy, B. R., Markoff, L. J., Chanock, R. M., Spring, S. B., Maassab, H. F., Kendal, A. P., Cox, N. J., Levine, M. M., Douglas, R. G., jr., Betts, R. F., Couch, R. B., Cate, T. R., jr.: Genetic approaches to attenzation of influenza A viruses for man. Phil. Trans. R. Soc. (Lond.) *B 288,* 401–415 (1980 a).

Murphy, B. R., Tolpin, M. D., Massicot, J. G., Kim, H. Y., Parrott, R. M., Chanock, R. M.: Escape of a highly defective influenza A virus mutant from its temperature sensitive phenotype by extragenic suppression and other types of mutation. Ann. N. Y. Acad. Sci. *345,* 172–182 (1980 b).

Nakajima, K., Sugiura, A.: Isolation of an influenza virus temperature-sensitive mutant of a new recombination-complementation group. Virology *78,* 365–374 (1977).

Nayak, D. P.: Defective interfering influenza viruses. Ann. Rev. Microbiol. *34,* 619–644 (1980).

Nayak, D. P., Tobita, K., Janda, J. M., Davis, A. R., De, B. K.: Homologous interference mediated by defective interfering influenza virus derived from a temperature-sensitive mutant of influenza virus. J. Virol. *28,* 375–386 (1978).

Nichol, S. T.: *In vitro* transcription of fowl plague virus RNA. Ph. D. thesis, University of Cambridge, 1980.

Nichol. S. T., Penn, C. R., Mahy, B. W. J.: Evidence for the involvement of influenza A (fowl plague Rostock) virus protein P2 in ApG and mRNA primed *in vitro* RNA synthesis. J. Gen. Virol. *57,* 407–413 (1981).

Odagiri, T., DeBorde, D. C., Maassab, H. F.: Cold-adapted recombinants of influenza A virus in MDCK cells. 1. Development and characterization of A/Ann Arbor/6/60 × A/Alaska/6/77 recombinant viruses. Virology *119,* 82–95 (1982).

Okhuma, S., Poole, B.: Fluorescence probe measurement of the intralysosomal pH in living cells and the perturbation of pH by various agents. Proc. Nat. Acad. Sci. *75,* 3327–3331 (1978).

Oxford, J. S.: Specific inhibitors of influenza virus replication as potential chemoprophylactic agents. J. Antimicrobial Chemotherapy *1,* 7–23 (1975).

Oxford, J. S., Corchoran, T., Schild, G. C.: Naturally occurring temperature-sensitive influenza A viruses of the H1N1 and H3N2 subtypes. J. Gen. Virol. *48,* 383–389 (1980).

Oxford, J. S., Logan, I. S., Potter, C. W.: Passage of influenza strains in the presence of aminoadamantane. Ann. N. Y. Acad. Sci. *173,* 300–313 (1970 a).

Oxford, J. S., Logan, I. S., Potter, C. W.: *In vivo* selection of an influenza A2 strain resistant to amantadine. Nature *226,* 82–83 (1970 b).

Oxford, J. S., Williams, J. D. (eds): Chemotherapy and Control of Influenza. London: Academic Press 1976.

Palese, P.: The genes of influenza virus. Cell *10,* 1–10 (1977).

Palese, P.: The hemagglutinin gene of influenza viruses. In: Topics in Infectious Diseases, Vol. 3: Influenza Virus Hemagglutinin (Laver, W. G., Bachmayer, H., Weil., R., eds.), 49–57. Wien-New York: Springer 1978.

Palese, P., Racaniello, V. R., Desselberger, U., Young, J., Baez, M.: Genetic structure and genetic variation of influenza viruses. Phil. Trans. R. Soc. (Lond.) *B 288,* 299–305 (1980).

Palese, P., Ritchey, M. B.: Live attenuated influenza virus vaccines: strains with temperature sensitive defects in P3 protein and nucleoprotein. Virology *78,* 183–191 (1977).

Palese, P., Ritchey, M. B., Schulman, J. L.: Mapping of the influenza virus genome. II. Identification of the P1, P2, and P3 genes. Virology *76,* 114–121 (1977 a).

Palese, P., Ritchey, M. B., Schulman, J. L.: P1 and P3 proteins of influenza virus are required for complementary RNA synthesis. J. Virol. *21,* 1187–1195 (1977 b).

Palese, P., Schulman, J. L.: Differences in RNA patterns of influenza A viruses. J. Virol. *17,* 876–884 (1976 a).

Palese, P., Schulman, J. L.: Mapping of the influenza virus genome: identification of the hemagglutinin and the neuraminidase genes. Proc. Natl. Acad. Sci. U.S.A. *73,* 2142–2146 (1976 b).

Palese, P., Tobita, K., Ueda, M., Compans, R. W.: Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. Virology *61,* 397–410 (1974).

Palese, P., Young, J. F.: Variation of influenza A, B, and C viruses. Science *215,* 1468–1474 (1982).

Penn, C. R.: Influenza virus messenger RNA synthesis. Ph. D. Thesis, University of Cambridge, 1983.

Penn, C. R., Blaas, D., Kuechler, E., Mahy, B. W. J.: Identification of the cap-binding protein of two strains of influenza A/FPV. J. Gen. Virol. *62,* 177–180 (1982).

Pereira, M. S.: Persistence of influenza in a population. In: Virus Persistence (Mahy, B. W. J., Minson, A. C., Darby, G. K., eds.), 15–37. (Society for General Microbiology Symposium, 33.) Cambridge University Press 1982.

Petri, T., Patterson, S., Dimmock, N. J.: Polymorphism of the NS1 proteins of type A influenza virus. J. Gen. Virol. *61,* 217–231 (1982).

Pons, M. W.: On the nature of the influenza virus genome. In: Current Topics Microbiol. Immunol. *52,* 142–157 (1970).

Pons, M. W.: The inhibition of influenza virus RNA synthesis by actinomycin D and cycloheximide. Virology *51,* 120–128 (1973).

Pons, M. W.: A re-examination of influenza single and double stranded RNAs by gel electrophoresis. Virology *69,* 789–792 (1976).

Pons, M. W.: Genome of incomplete influenza virus. Virology *100,* 43–52 (1980).

Porter, A. G., Barber, C., Carey, N. H., Hallewell, R. A., Threlfall, G., Emtage, J. S.: Complete nucleotide sequence of an influenza virus haemagglutinin gene from cloned DNA. Nature *282,* 471–477 (1979).

Porter, A. G., Smith, J. C., Emtage, J. S.: Nucleotide sequence of influenza virus RNA segment 8 indicates that coding regions for NS_1 and NS_2 overlap. Proc. Natl. Acad. Sci. U.S.A. *77,* 5074–5078 (1980).

Portner, A., Webster, R. G., Bean, W. H.: Similar frequencies of antigenic variation in Sendai, vesicular stomatitis and influenza A viruses. Virology *104,* 235–238 (1980).

Pringle, C. R.: The genetics of vesiculoviruses. Arch. Virol. *72,* 1–34 (1982).

Pringle, C. R., Devine, V., Wilkie, M., Preston, C. M., Dolan, A., McGeoch, D. J.: Enhanced mutability associated with a temperature sensitive mutant of vesicular stomatitis virus. J. Virol. *39,* 377–388 (1981).

Ramig, R. F., Fields, B. N.: Revertants of temperature-sensitive mutants of reovirus: evidence for frequent extragenic suppression. Virology *92,* 155–167 (1979).

Richman, D. D., Yazaki, P., Hostetler, K. Y.: The intracellular distribution and antiviral activity of amantadine. Virology *112,* 81–90 (1981).

Ritchey, M. B., Palese, P.: Identification of the defective genes in three mutant groups of influenza virus. J. Virol. *21,* 1196–1204 (1977).

Ritchey, M. B., Palese, P., Kilbourne, E. D.: RNAs of influenza A, B, and C viruses. J. Virol. *18,* 738–744 (1976 a).

Ritchey, M. B., Palese, P., Schulman, J. L.: Mapping of the influenza virus genome. III. Identification of genes coding for nucleoprotein, membrane protein and nonstructural protein. J. Virol. *20,* 307–313 (1976 b).

Ritchey, M. B., Palese, P., Schulman, J. L.: Differences in protein patterns of influenza A viruses. Virology *76,* 122–128 (1977).

Ritchie, D. A.: Genetic analysis of animal viruses. Brit. Med. Bull. *29,* 247–252 (1973).
Robertson, J. S.: 5′ and 3′ Terminal nucleotide sequences of the RNA genome segments of influenza virus. Nucleic Acids Res. *6,* 3745–3757 (1979).
Robertson, J. S., Robertson, E., Roditi, I., Almond, J. W., Inglis, S. C.: Sequence analysis of fowl plague virus mutant *ts*47 reveals a nonsense mutation in the NS_1 gene. Virology (in press, 1983).
Robertson, J. S., Schubert, M., Lazzarini, R. A.: Polyadenylation sites for influenza virus mRNA. J. Virol. 38, 157–163 (1981).
Rochovansky, O. M.: RNA synthesis by ribonucleoprotein-polymerase complexes isolated from influenza virus. Virology *73,* 327–338 (1976).
Rohde, W., Harms, E., Scholtissek, C.: Biochemical studies on influenza viruses. I. Comparative analysis of equine 2 virus and virus N genes and gene products. Virology *79,* 393–404 (1977).
Rohde, W., Harms, E., Scholtissek, C.: Studies on the genome structure of influenza A viruses. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 11–17. London: Academic Press 1978.
Rohde, W., Boschek, C. B., Harms, E., Rott, R., Scholtissek, C.: Characterization of virus-like particles produced by an influenza A virus. Arch. Virol. *62,* 291–302 (1979).
Rohde, W., Scholtissek, C.: On the origin of the gene coding for an influenza A virus nucleocapsid protein. Arch. Virol. *64,* 213–223 (1980).
Rott, R.: Molecular basis of infectivity and pathogenicity of myxovirus. Arch. Virol. *59,* 285–298 (1979).
Rott, R., Orlich, M., Scholtissek, C.: Correlation of pathogenicity and gene constellation of influenza A viruses. III. Non-pathogenic recombinants derived from highly pathogenic parent strains. J. Gen. Virol. *44,* 471–477 (1979).
Rott, R., Orlich, M., Scholtissek, C.: Differences in the multiplication at elevated temperature of influenza virus recombinants pathogenic and non-pathogenic for chicken. Virology *120,* 215–224 (1982).
Scholtissek, C.: Synthesis *in vitro* of RNA complementary to parental viral RNA by RNA polymerase induced by influenza virus. Biochim. Biophys. Acta *179,* 389–397 (1969).
Scholtissek, C.: The genome of the influenza virus. In: Current Topics in Microbiology and Immunology, Vol. 80, 139–169, 1978.
Scholtissek, C.: Influenza virus genetics. Adv. Genetics *20,* 1–36 (1979).
Scholtissek, C., Bowles, A. L.: Isolation and characterization of temperature-sensitive mutants of fowl plague virus. Virology *67,* 576–587 (1975).
Scholtissek, C., Faulkner, G. P.: Amantadine-resistant and -sensitive influenza strains and recombinants. J. Gen. Virol. *44,* 807–815 (1979).
Scholtissek, C., Harms, E., Rohde, W., Orlich, M., Rott, R.: Correlation between RNA fragments of fowl plague virus and their corresponding gene functions. Virology *74,* 332–344 (1976).
Scholtissek, C., Koennecke, I., Rott, R.: Host range recombinants of fowl plague (influenza A) virus. Virology *91,* 79–85 (1978 a).
Scholtissek, C., Kruczinna, R., Rott, R., Klenk, H.-D.: Characteristics of an influenza mutant temperature-sensitive for viral RNA synthesis. Virology *58,* 317–322 (1974).
Scholtissek, C., Murphy, B. R.: Host range mutants of an influenza A virus. Arch. Virol. *58,* 323–333 (1978).
Scholtissek, C., Rohde, W., Harms, E., Rott, R., Orlich, M., Boschek, C. B.: A possible partial heterozygote of an influenza A virus. Virology *89,* 506–516 (1978 b).
Scholtissek, C., Rott, R.: Behavior of virus-specific activities in tissue cultures infected with myxoviruses after chemical changes of the viral ribonucleic acid. Virology *22,* 169–176 (1964).
Scholtissek, C., Rott, R.: Effect of temperature on the multiplication of an influenza virus. J. Gen. Virol. *5,* 283–290 (1969 b).
Scholtissek, C., Rott, R.: Ribonucleic acid nucleotidyl transferase induced in chick fibroblasts after infection with an influenza virus. J. Gen. Virol. *4,* 125–137 (1969 a).
Scholtissek, C., Rott, R.: Synthesis *in vivo* of influenza virus plus and minus strand RNA and its preferential inhibition by antibiotics. Virology *40,* 989–996 (1970).
Scholtissek, C., Spring, S. B.: Suppressor recombinants of an influenza A virus. In: The Replication of Negative Strand Viruses (Bishop, D. H. L., Compans, R. W., eds.), 389–394 (1981 a).

Scholtissek, C., Spring, S. B.: Suppressor recombinants and suppressor mutants. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., ed.), 399–413. Academic Press 1981 b.

Scholtissek, C., Spring, S. B.: Extragenic suppression of temperature-sensitive mutations in RNA segment 8 by replacement of different RNA segments with those of other influenza A virus prototype strains. Virology *118,* 28–34 (1982).

Schulman, J. L., Palese, P.: Selection and identification of influenza virus recombinants of defined genetic composition. J. Virol. *20,* 248–254 (1976).

Schulman, J. L., Palese, P.: Virulence factors of influenza A viruses: WSN virus neuraminidase required for plaque production in MDBK cells. J. Virol. *24,* 170–176 (1977).

Schuster, H., Schramm, G.: Bestimmung der biologisch wirksamen Einheit in der Ribonucleinsäure des Tabakmosaikvirus auf chemischem Wege. Z. Naturforsch. *13 b,* 697–704 (1958).

Schwartz, D. O., Beckwith, J. R.: Mutagens which cause deletions in E. coli. Genetics *61,* 371–376 (1969).

Schwartz, R. T., Scholtissek, C.: Purification and properties of the RNA polymerase-template complex of an influenza virus. Z. Naturforsch. *28 C,* 202–207 (1973).

Shaw, M. W., Compans R. W.: Isolation and characterization of cytoplasmic inclusions from influenza A virus-infected cells. J. Virol. *25,* 605–615 (1978).

Shimizu, K., Murphy, B. R., Chanock, R. M.: Temperature-sensitive mutants of influenza A/Udorn/72(H3N2) virus: intrasegmental complementation and temperature-dependent host range *(td-hr)* mutation. In: The Replication of Negative Strand Viruses (Bishop, D. H. L., Compans, R. W., eds.), 370–378. Elsevier/North-Holland 1981.

Shimizu, K., Mullinix, M. G., Chanock, R. M., Murphy, B. R.: Temperature-sensitive mutants of influenza A/Udorn/72 (H3N2) virus. I. Isolation of temperature-sensitive mutants some of which exhibit host-dependent temperature sensitivity. Virology *117,* 38–44 (1982 a).

Shimizu, K., Mullinix, M. G., Chanock, R. M., Murphy, B. R.: Temperature sensitive mutants of influenza A/Udorn/72 (H3N2) virus. II. Genetic analysis and demonstration of intrasegmental complementation. Virology *117,* 45–61 (1982 b).

Shimizu, K., Mullinix, M. G., Chanock, R. M., Murphy, B. R.: Temperature-sensitive mutants of influenza A/Udorn/72 (H3N2) virus. III. Genetic analysis of temperature-dependent host range mutants. Virology *124,* 35–44 (1983).

Simpson, R. W., Bean, W. J.: The biologically active proteins of influenza virus. Influenza transcriptase activity of cells and virions. In: The Influenza Virus and Influenza (Kilbourne, E. D., ed.), 125–143. New York: Academic Press 1975.

Simpson, R. W., Hirst, G. K.: Genetic recombination among influenza viruses. I. Cross reaction of plaque-forming capacity as a method for selecting recombinants from the progeny of crosses between influenza A strains. Virology *15,* 436–451 (1961).

Simpson, R. W., Hirst, G. K.: Temperature-sensitive mutants of influenza A virus: isolation of mutants and preliminary observations on genetic recombination and complementation. Virology *35,* 41–49 (1968).

Skehel, J. J., Burke, D. C.: Ribonucleic acid synthesis in chick embryo cells infected with fowl plague virus. J. Virol. *3,* 429–438 (1969).

Skehel, J. J., Hay, A. J.: Nucleotide sequences at the 5′-termini of influenza virus RNAs and their transcripts. Nucleic Acids Res. *5,* 1207–1219 (1978 a).

Skehel, J. J., Hay, A. J.: Influenza virus transcription. J. Gen. Virol. *39,* 1–8 (1978 b).

Skehel, J. J., Hay, A. J., Armstrong, J. A.: On the mechanism of inhibition of influenza virus replication by amantadine hydrochloride. J. Gen. Virol. *38,* 97–110 (1977).

Smith, G. L., Hay, A. J.: Replication of the influenza virus genome. Virology *118,* 96–108 (1982).

Spindler, K. R., Horodyski, F. M., Holland, J. J.: High multiplicities of infection favor rapid and random evolution of vesicular stomatitis virus. Virology *119,* 96–108 (1982).

Spring, S. B., Nusinoff, S. R., Mills, J., Richman, D. D., Tierney, E. L., Murphy, B. R., Chanock, R. M.: Temperature-sensitive mutants of influenza virus. VI. Transfer of *ts* lesions from the Asian subtype of influenza A virus (H2N2) to the Hong Kong subtype (H3N2). Virology *66,* 522–532 (1975 a).

Spring, S. B., Nusinoff, S. R., Tierney, E. L., Richman, D. D., Murphy, B. R., Chanock, R. M.: Temperature-sensitive mutants of influenza. VIII. Genetic and biological characterization of *ts* mutants of

influenza virus A (H3N2) and their assignment to complementation groups. Virology *66*, 542–550 (1975 b).

Spring, S. B., Maassab, H. F., Kendal, A. P., Murphy, B. R., Chanock, R. M.: Cold-adapted variants of influenza virus A. 1. Comparison of the genetic properties of *ts* mutants and five cold-adapted variants of influenza virus A. Virology *77*, 337–343 (1977 a).

Spring, S. B., Maassab, H. F., Kendal, A. P., Murphy, B. R., Chanock, R. M.: Cold-adapted variants of influenza A. II. Comparison of the genetic and biological properties of *ts* mutants and recombinants of the cold-adapted A/AA/6/60 strain. Arch. Virol. *55*, 233–246 (1977 b).

Staroff, A., Bukrinskaya, A.: Different mechanisms of synthesis of complete and incomplete influenza virus transcripts in infected cells. Arch. Virol. *67*, 105–110 (1981).

Sugiura, A.: Influenza virus genetics. In: The Influenza Viruses and Influenza (Kilbourne, E. D., ed.), 171–213. New York: Academic Press 1975.

Sugiura, A., Tobita, K., Kilbourne, E. D.: Isolation and preliminary characterization of temperature-sensitive mutants of influenza virus. J. Virol. *10*, 639–647 (1972).

Sugiura, A., Ueda, M., Tobita, K., Enomoto, C.: Further isolation and characterization of temperature-sensitive mutants of influenza virus. Virology *65*, 363–373 (1975).

Taylor, J. M., Illmensee, R., Litwin, S., Herring, L., Broni, B., Krug, R. M.: Use of specific radioactive probes to study transcription and replication of the influenza virus genome. J. Virol. *21*, 530–540 (1977).

Thierry, F., Danos, O.: Use of specific single stranded DNA probes cloned in M13 to study the RNA synthesis of four temperature-sensitive mutants of HK/68 influenza virus. Nucleic Acids Res. *10*, 2925–2937 (1982).

Thierry, F., Spring, S. B.: Localization of the *ts* defects of *ts* mutants of influenza A virus using complementation analysis and gel analysis of the RNA segments of recombinants. Virology *115*, 137–148 (1981).

Thierry, F., Spring, S. B., Chanock, R. M.: Localization of the *ts* defect in two *ts* mutants of influenza A virus: evidence for the occurrence of intracistronic complementation between *ts* mutants of influenza A virus coding for the neuraminidase and nucleoprotein polypeptides. Virology *101*, 484–492 (1980).

Tolpin, M. D., Massicot, J. G., Mullinix, M. G., Kim, H. W., Parrott, R. H., Chanock, R. M., Murphy, B. R.: Genetic factors associated with loss of the temperature-sensitive phenotype of the influenza A/Alaska/77-*ts*-1A2 recombinant during growth *in vivo*. Virology *112*, 505–517 (1981).

Tuckova, E., Vonka, V.: Genetic interaction between two influenza A/WS virus mutants. III. Capacity of WS/NWS recombinants to reproduce in various cell systems. Acta Virol. *17*, 501–504 (1973).

Tuckova, E., Vonka, V., Starek, M.: Genetic interaction between two influenza A/WS virus strains. 1. Characterization of the mutants. Acta Virol. *15*, 337–344 (1971).

Tuckova, E., Vonka, V., Zavadova, A., Kutineva, L.: Sensitivity to 1-adamantanamine as a marker in genetic studies with influenza viruses. J. Biol. Stand. *1*, 341–346 (1973).

Ueda, M.: Temperature-sensitive mutants of influenza virus. Isolation and preliminary characterization. Arch. Virusforsch. *39*, 360–368 (1972).

Ueda, M., Kilbourne, E. D.: Temperature-sensitive mutants of influenza virus: a mutation in the hemagglutinin gene. Virology *70*, 425–431 (1976).

Ulmanen, I., Broni, B. A., Krug, R. M.: Role of two of the influenza virus core P proteins in recognizing cap 1 structures (m^7GpppNm) on RNAs and in initiating viral RNA transcription. Proc. Natl. Acad. Sci. U.S.A. *78*, 7355–7359 (1981).

Ulmanen, I., Broni, B., Krug, R. M.: Influenza virus temperature-sensitive cap (m^7GpppNm)-dependent endonuclease. J. Virol. *45*, 27–35 (1983).

Von Magnus, P.: Incomplete forms of influenza virus. Adv. Virus Res. *2*, 59–79 (1954).

Webster, R. G., Laver, W. G., Air, G. M., Schild, G. C.: Molecular mechanisms of variation in influenza viruses. Nature *296*, 115–121 (1982).

Wilson, I. A., Skehel, J. J., Wiley, D. C.: Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature *289*, 366–373 (1981).

Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M., Smith M.: Redesigning enzyme structure by site-directed mutagenesis: tyrosyl tRNA synthetase and ATP binding. Nature *299*, 756–758 (1982).

Winter, G., Fields, S.: 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 (1980).

Winter, G., Fields, S.: The structure of the gene encoding the nucleoprotein of human influenza virus A/PR/8/34. Virology *114*, 423–428 (1981).

Winter, G., Fields, S.: Nucleotide sequence of human influenza A/PR/8/34 segment 2. Nucleic Acids Res. *10*, 2135–2143 (1982).

Winter, G., Fields, S., Brownlee, G. G.: Nucleotide sequence of the haemagglutinin gene of a human influenza virus H1 subtype. Nature *292*, 72–75 (1981 a).

Winter, G., Fields, S., Gait, M. J., Brownlee, G. G.: The use of synthetic oligodeoxynucleotide primers in cloning and sequencing segment 8 of influenza virus A/PR/8/34. Nucleic Acids Res. *9*, 237–245 (1981 b).

Winter, G., Fields, S., Pratti, G.: The structure of two subgenomic RNAs from human influenza A/PR/8/34. Nucleic Acids Res. *9*, 6907–6915 (1981 c).

Wolstenholme, A. J., Barrett, T., Nichol, S. T., Mahy, B. W. J.: Influenza virus-specific RNA and protein syntheses in cells infected with temperature-sensitive mutants defective in the genome segment encoding nonstructural proteins. J. Virol. *35*, 1–7 (1980).

Yoshida, T., Shaw, M. W., Young, J. F., Compans, R. W.: Characterization of the RNA associated with influenza A cytoplasmic inclusions and the interaction of NS_1 protein with RNA. Virology *110*, 87–97 (1981).

Zavada, J.: A host-cell DNA function involved in the replication of avian tumour viruses and of fowl-plague virus. J. Gen. Virol. *4*, 571–576 (1969).

Zhang, Y., Tian, S., Zhu, J.: Identification of naturally occurring temperature-sensitive strains of influenza A virus and location of their genetic lesions. Scientia Sinica *B 25*, 411–419 (1982).

Zvonarjev, A. Y., Ghendon, Y. Z.: Influence of membrane (M) protein of the influenza A virus virion transcriptase activity *in vitro* and its susceptibility to rimantadine. J. Virol. *33*, 583–586 (1980).

8

The Structure of Influenza Virus Defective Interfering (DI) RNAs and Their Progenitor Genes

D. P. Nayak and *N. Sivasubramanian*

I. Introduction

The term defective interfering (DI) virus particles was used by Huang and Baltimore [31] to define a class of virus particles which possess the following properties: (a) They are defective, unable to replicate independently; (b) They require the helper function of standard virus for replication; (c) They interfere with the replication of standard virus; (d) They are usually deletion mutants, *i.e.*, the genome of DI particles does not represent the entire genome of standard virions. Although all of the DI genomes studied to date are shorter than the standard viral genome and are deletion mutants, it is possible some DI genome may contain an altered nucleic acid sequence(s) rather than a deletion(s) and become defective as well as interfering. DI particles were first observed by von Magnus [64,65] when he passaged influenza virus serially undiluted in embryonated chicken eggs. He called these particles "incomplete" particles which were both defective and interfering. Subsequently, DI particles have been reported for nearly all animal viruses during high multiplicity passages (for a detailed review see reference no. [32]).

Although influenza virus DI particles were first demonstrated over 30 years ago [64, 65], only recently have their properties been studied in detail. Some observations on the properties of influenza virus DI particles and their genomes were reviewed earlier [48, 50]. Since then, complete sequences of the progenitor genes (PB1, PA and PB2) and a number of DI RNAs have been determined using cloning and DNA sequencing methods. Therefore, in this chapter we shall place a major emphasis on reviewing the newer information on the structure of the DI RNAs and their relationships to the progenitor genes.

II. Properties of Influenza DI Particles

A number of workers [14, 32, 47] have confirmed von Magnus's [64, 65] observations on the formation of noninfectious virus by passaging undiluted virus in either

embryonated chicken eggs [47] or cell culture. Using WSN virus and MDBK cell systems, DI particles could be produced efficiently when passaged at high multiplicity. By passage 3 to passage 5, DI particles replaced over 99% of the standard virus particles. These DI particles could be amplified and maintained (at least for a number of passages) using the standard virus as a helper. The interfering property could be quantified accurately as DIU (defective interfering units)/ml using a plaque inhibition assay [12, 14, 35]. It was further shown that a single DI particle is capable of inhibiting plaque formation by standard virus particles. The inhibition of visible plaques by standard virus particles was due to the reduction of the cytopathic effect by DI particles in coinfected cells [35].

Although influenza virus DI particle preparations contain aberrant particles [66] they possess essentially the same density and proteins as standard virus. Only minor differences in the quantity of NP, M, or glycoproteins (HA and NA) have been observed [42, 46].

Influenza virus DI particles are rather unstable; the population of DI particles during prolonged passages often undergoes further evolution with the apparent emergence of new DI particles replacing the previous ones. Since DI particles suppress cytopathic effects, they aid in establishing persistent infections of MDBK and HeLa cells *in vitro* with influenza viruses [25].

III. RNA of Influenza Virus DI Particles

It was reported earlier that DI particles contain less RNA and less RNP compared to the standard virus [1, 2, 42]. In addition, RNA from DI particles was found to contain consistently one or more smaller RNA segments (DI RNA) [20, 21, 44, 49, 61] not present in standard virus (Fig. 1). Furthermore, as reported earlier, there was a greater reduction (or even a complete loss) of the larger RNA segments [13, 18, 20, 22, 61]. However, since the decrease of a specific RNA segment(s) in a DI virus preparation could not often account for the total loss of infectivity [5], it was postulated [48] that the loss of infectivity was due to a random rather than a specific loss of one or more RNA segments and the novel RNA segments found in DI particles were thought to be the interfering molecules [34]. Subsequently, partial complementation [37] at a higher multiplicity of DI particles has supported this hypothesis that the greatly reduced infectivity in DI particles is due to a random rather than a specific loss of one or more viral RNA segments and that full complementation at a higher multiplicity of DI particles is prevented by the presence of DI RNAs in the DI particles. Several lines of evidence suggest that these small RNA (DI RNA) molecules are responsible for the interference: (i) DI RNAs are consistently present in DI preparations and are absent in repeatedly cloned standard virus preparations; (ii) These small RNA molecules are formed and become amplified during repeated high multiplicity passages; their presence and quantity often correlate with the interfering property (DIU/ml) of a preparation; (iii) Ultraviolet light sensitivity data suggest that, the target size of the interfering mulecule is small and approximates the average size of the small RNA species in the DI preparation; (iv) Finally, ribonucleoprotein (RNP) complexes containing DI

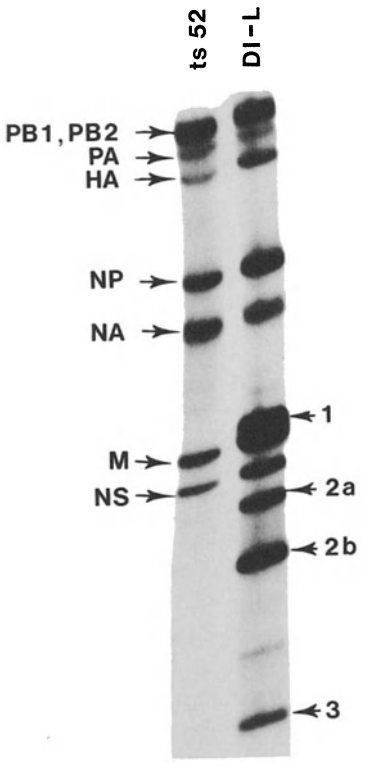


Fig. 1. PAGE analysis of standard and DI RNA (L) segments [23]. Labeled RNA was isolated from purified standard and DI virus particles and analyzed by electrophoresis in a composite gel (polyacrylamide, 2.2%, agarose, 0.6% and urea, 6M) for 23 hours at 180 V and 4 °C [23]

RNA molecules were interfering, whereas various RNP complexes containing standard RNA segments were not interfering [34]. Therefore, influenza DI particles appear to contain less than a full complement of standard viral genes because of the loss of one or more RNA segments and, in addition, they contain one or more DI RNA segments which are responsible for interference.

These DI RNAs, when further analyzed, were found to be of the same polarity as viral RNA segments and, unlike the majority the DI RNAs of the non-segmented negative strand virus, vesicular stomatitis virus, did not possess complementary termini. Hybridization, oligonucleotide mapping and RNA sequencing of the 5′ and 3′ termini further showed that all of the 16 DI RNAs studied to date are of polymerase gene (PB1, PB2, PA) origin [21, 23, 24, 44] and that they contain both the 5′ and 3′ genomic termini [23, 51]. Oligonucleotide mapping showed that DI RNAs of different size can be generated from a single polymerase gene and that the sequences of smaller DI RNAs are not always subsets of the larger DI RNAs. It was, therefore, postulated that at least some of the DI RNAs are formed from internal deletions of a polymerase gene with the preservation of both termini [23]. However, these studies could not reveal the exact sequence relationships among

Table 1. *Nucleotide frequencies in influenza virus genes (vRNA)*

	PB1	PA	PB2	HA	NP	NA	M	NS	Average
U	34.4	33.3	33.6	34.8	32.5	29.6	28.6	33.3	32.5
G	19.9	18.8	18.4	18.3	20.1	18.4	21.1	19.8	19.4
C	22.4	23.6	25.5	22.8	25.9	25.3	26.1	23.4	24.3
A	23.3	24.2	22.5	24.2	21.5	26.8	24.1	23.5	23.8

The nucleotide frequences of influenza virus genes were calculated from the published nucleotide sequences of individual genes. The sequences of PB1, PB2, HA and NA genes of A/WSN/33 [36, 58, 29 a, 29 b], NP and M genes of A/PR/8/34 [67,68], NS gene of A/Udorn/72 [39 a] and PA gene of A/NT/60/68 [8] were used. Nucleotide frequencies = (Number of specific nucleotide/total number of nucleotides in a specific gene) × 100.

the DI RNAs or between the progenitor gene and its DI RNAs, since such information can be obtained only from detailed sequence analyses. Furthermore, since polymerase proteins are presumed to play an important role in the generation of these DI RNAs and in DI-mediated interference (see sections 6 and 7) better understanding of the primary and the secondary structures of these polymerase proteins should illuminate the details of the RNA-protein interactions necessary for both the transcription and replication of standard RNAs and the generation of DI RNAs as well. Therefore, a comparative analysis of the nucleotide sequences of the three polymerase genes and of the predicted primary and secondary structures of PB1 and PB2 proteins will be presented in the following section (section 4). The secondary structure analysis of PA protein is in progress.

IV. Structure of Polymerase Genes and Polymerase Proteins

A. Nucleotide Sequence of Polymerase Genes

The three polymerase genes account for over half of the genetic information in the influenza virus genome. With the aid of recombinant DNA technology, the complete sequences of all three polymerase genes of influenza A viruses have been determined [7, 8, 27, 36, 58, 69]. The PB1 gene of WSN, A/NT/60/68 and PR/8 is 2341 nucleotides long with the first AUG at position 25 and an open reading frame of 2271 nucleotides (positions 25—2295) encoding 757 amino acids. The encoded PB1 protein is basic, with a molecular weight of 86,500. The PA gene of PR/8 and A/NT/60/68 consists of 2233 nucleotides with a single reading frame encoding 716 amino acids and the first AUG at the position 25. The encoded protein is slightly acidic, with a molecular weight of 82,400. The PB2 gene of both WSN and PR/8 consists of 2341 nucleotides encoding a basic protein of 759 amino acids with the first AUG at position 27. Table 1 shows the frequency of nucleotides in influenza virus genes.

B. Primary Structure of PB1 and PB2 Proteins

The primary structure of PB1 and PB2 proteins of WSN and PR/8 viruses as predicted from nucleotide sequences has been reported [27, 36, 58, 69]. Since both PB1 and PB2 genes possess an identical number of nucleotides and encode basic proteins of similar sizes and since both proteins are also involved in the transcription process and may interact with the same viral RNA templates, a comparison of the primary and secondary structures of these two proteins may reveal some common features with respect to modes of RNA-protein interaction. To determine if any homology existed at the level of primary structure, both inter-and intra-protein homology and the corresponding nucleotide sequence homologies were determined (Table 2). A striking feature is an iterative tetrapeptide beginning at amino acid residue 722 (Ala-Arg-Ile-Asp-Ala-Arg-Ile-Asp) in the PB1 protein of WSN, A/NT/60/68 and PR/8 viruses [7, 58, 69]. The RNA which codes for this region is equally iterative, with only a single nucleotide mismatch in the WSN PB1 gene [58]. The chance occurrence of such an iterative tetrapeptide and nucleotide sequence is rare and suggestive of a duplication event in the evolution of the PB1 gene. This octapeptide is predicted to form an α-helix (Fig. 2).

The amino terminus of the PB1 protein contains fewer charged residues (3 out of 31) than that of the PB2 protein (11 out of 31). The carboxy terminus of the PB1 protein contains more charged residues (18 out of 37) than the PB2 protein (9 out of 37). In addition, the PB1 amino terminus is also more hydrophobic (16 out of 31) than the PB2 (12 out of 31) amino terminus. Charge calculations indicate that the PB1 protein is more basic than the NP and the M proteins but slightly less basic than the PB2 protein [36].

The PB1 protein contains many clusters of basic amino acids in regions predicted to be devoid of secondary structure (*e.g.*, amino acid residues starting from 187, 207, 429, and 479; ref. [58]). These clusters contain 3 to 4 arginine and lysine residues in close proximity without being interrupted by acidic residues. Such clusters of basic amino acids are similar to that present in the PB2 protein but are much more pronounced than those reported for the PR/8 NP [62, 68] and M [67] proteins. Clusters of basic amino acids are also found in the α-helical regions of the protein molecule. Since the PB1 protein is likely to interact with the template viral

Table 2. *Nucleotide homologies with the corresponding peptide homologies of PB1 and PB2 proteins*

Gene	Nucleotide number	Nucleotide homology	Protein	Amino acid number	Amino acid homology
PB1	2188	GCCCGAAUUGAU	PB1	722	Ala Arg Ile Asp
PB1	2200	GCACGAAUUGAU	PB1	726	Ala Arg Ile Asp
PB1	2080	GAACAAAUGUAC	PB1	686	Glu Gln Met Tyr
PB2	748	GAACAGAUGUAC	PB2	241	Glu Gln Met Tyr
PB1	355	AUGGAGGUUGUU	PB1	111	Met Glu Val Val
PB2	517	AUGGAAGUUGUU	PB2	164	Met Glu Val Val
PB1	649	AAGCAGAGAUUGAAC	PB1	210	Lys Gln Arg Leu Asn
PB2	1300	AAUCAGCGAUUGAAC	PB2	426	Asn Gln Arg Leu Asn

Underlines show mismatches in homologous regions.

RNA during the initiation of transcription, these clusters of basic residues might play a vital role in catalyzing this process. Similar RNA-protein interactions via clusters of basic amino acids have been proposed for influenza PB2 [36], influenza NP [68], Semliki forest virus nucleocapsid [29], VP1 of SV40 [63] and polyoma virus [60], and the core antigen of hepatitis virus [52].

Table 2 also shows the occurrence of a few homologies between the nucleotide and amino acid sequences of PB1 and PB2. Since the homologies of the amino acid and nucleotide sequences are unlikely to occur by chance, they may be genetically significant. However, except for these limited similarities, the primary structures of both PB1 and PB2 genes and their proteins are different and suggest independent evolution of the multiple polymerase genes of influenza viruses.

A comparison of the amino acid groups among the basic proteins of influenza virus, MS2 replicase [28] and polio virus P3-1b [39] proteins revealed some interesting features. While all of the proteins which are involved in nucleic acid binding and synthesis are basic proteins, the content of arginine (the most preferred basic amino acid) is high among the basic proteins of influenza virus (Table 3). Since arginine residues in histones appear to play a vital role in organizing the eukaryotic nucleoprotein complex [10], they may also be involved in organizing the influenza nucleoprotein complex. It is also interesting to note that more hydrophobic amino acids are present in the PB2 protein (Table 3) and are distributed in longer stretches than in the PB1 protein. This might explain the proposed membrane attachment of the PB2 protein [6]. Also, all of the proteins (PB1, PB2, NP, MS2 replicase, polio

Table 3. *Representation of amino acid types in basic proteins of influenza virus (PB1 and PB2 from A/WSN/33 virus and NP, M from A/PR/8/34 virus) and other proteins having polymerase activities [16, 25]*

Amino acid types	Percent of each amino acid type						Average % in proteins[b]
	PB1	PB2	NP	M	MS2 Replicase	Polio P3-1b	
Small aliphatic (A, G)[a]	11.5	12.5	16.0	16.3	15.8	14.5	16.9
Hydroxyl (S, T)	14.7	14.1	13.6	14.2	14.5	12.5	13.1
Acidic (D, E)	10.5	10.7	9.9	9.2	10.1	11.6	11.6
Acidic + acid amide (D, N, E, Q)	21.3	20.0	22.1	19.6	16.2	19.2	19.8
Basic (K, R, H)	14.9	15.1	15.2	14.0	14.2	13.8	13.5
Hydrophobic (L, V, I, M)	23.3	27.1	22.6	26.7	22.6	24.6	20.2
Aromatic (F, Y, W)	8.7	6.6	6.6	5.2	10.7	9.5	8.3

[a] Abbreviations: A = Alanine, G = Glycine, S = Serine, T = Threonine, D = Glutamic acid, E = Aspartic acid, N = Asparagine, Q = Glutamine, K = Lysine, R = Arginine, H = Histidine, L = Leucine, V = Valine, I = Isoleucine, M = Methionine, F = Phenyl alanine, Y = Tyrosine, W = Tryptophan.

[b] The average amino acid composition of proteins is included for comparison. The average amino acid composition was calculated from the amino acid sequences of 314 proteins [24 a].

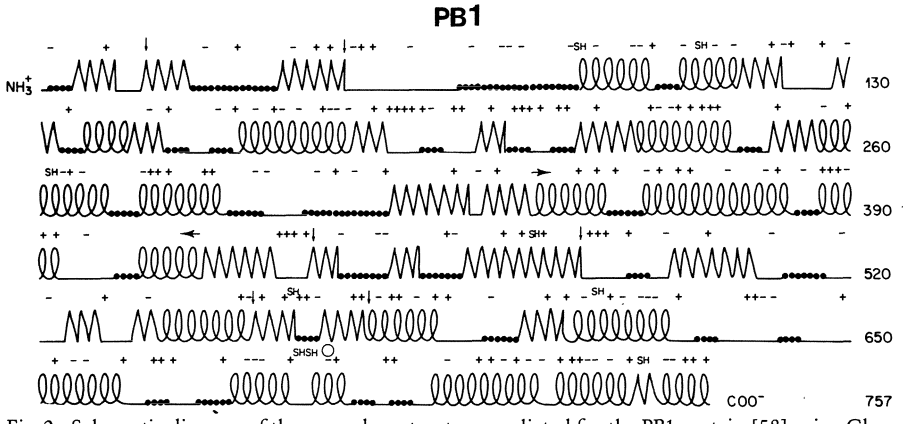


Fig. 2. Schematic diagram of the secondary structure predicted for the PB1 protein [58] using Chou-Fasman probablistic method [19 a]. ⌇⌇⌇⌇⌇⌇⌇⌇⌇⌇⌇ represents α-helix structures, ⩙⩙⩙⩙⩙⩙⩙ represents β-pleated sheets, ●●●●●● represents β-turns (chain reversals) and ———— represents random or undefined structure. +, — represent positive and negative charges and SH represents the location of cysteine residues. ↓↓ indicate the region of antiparallel β sheets, and → indicate the region of four helical supersecondary structure

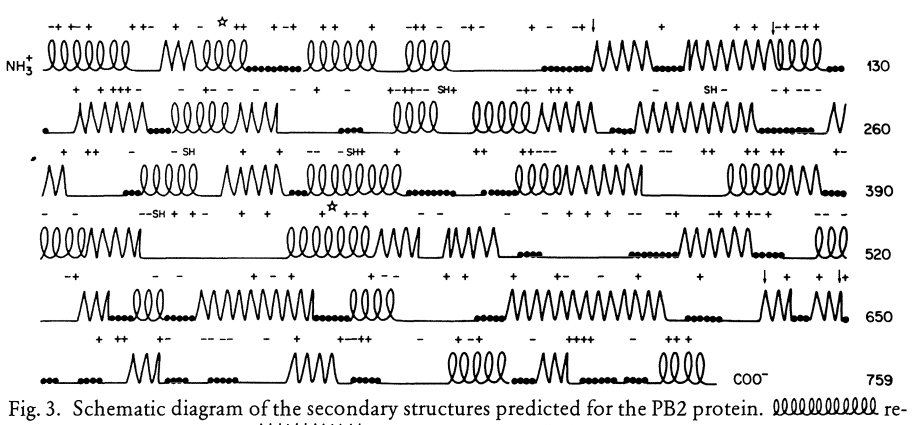


Fig. 3. Schematic diagram of the secondary structures predicted for the PB2 protein. ⌇⌇⌇⌇⌇⌇⌇⌇⌇⌇⌇ represents α-helix structure, ⩙⩙⩙⩙⩙⩙⩙ represents β-pleated sheets, ●●●●●● represents β-turns (chain reversals) and ———— represents random or undefined structure. +, — denotes the positive and negative charges and SH denotes the location of cysteine residues. ↓↓ indicate the region of antiparallel β-pleated sheets

P3-1b) involved in RNA polymerization possess a similar pattern in the composition of various amino acid groups (Table 3).

C. Predicted Secondary Structures of PB1 and PB2 Proteins

Since polymerase proteins are involved in complex biological processes such as transcription and replication, they must associate in protein-protein as well as RNA-protein interaction. An understanding of these processes would require the elucidation of the structure-function relationships of different domains within these proteins. Figures 2 and 3 illustrate the predicted secondary structure of PB1 and PB2 proteins and Table 4 shows the contents of the secondary structural elements of these proteins.

The PB1 protein contains a "super secondary structure" (four anti-parallel α-helices, amino acid residues 341–415, Fig. 4 A) which may be involved in the RNA-protein interaction process. Similar "four helical super secondary structures" are present in other proteins involved in either RNA or DNA interaction, *e.g.,* tobacco mosaic virus protein [16], tyrosyl-tRNA synthetase [33] and *Escherichia coli* DNA polymerase I [9]. The presence of many positive charges in these α-helices of the PB1 protein would also support their involvement in RNA binding. Also, the most hydrophilic (charged) α-helix (amino acid residues 386 to 393) and the most hydrophobic α-helix (amino acid residues 407 to 415, Fig. 4 A) as determined by helical hydrophobic moment plot analyses [58] of PB1 protein, constitute two of the α-helices of this super secondary structure. Such a structure is absent in PB2 protein. Another important structural feature is the presence of "polypeptide double helices" (antiparallel $\beta\beta$ dimer) in both proteins (Fig. 4, Band C). These polypeptide double helices have been proposed to be involved in the interaction with the minor groove of the RNA helix [11] and are found in DNA polymerase [9] and Lac repressor [19]. Four such structures are found in the PB1 protein (Fig. 4 B) conpared to two in the PB2 protein (Fig. 4 C). Additional helical hydrophobic moment analyses [58] show that none of the α-helices of the PB1 and PB2 proteins are of the transmembrane type. The majority of the α-helices of PB1 and PB2 proteins have medium hydrophobic moments and low mean hydrophobicities which are characteristic of soluble globular proteins. Exceptions are one α-helix in the PB1 protein (amino acid residues 695 to 700) and two α-helices in the PB2

Table 4. *Content of the predicted secondary structural elements in PB1 and PB2 proteins*

Secondary structure	PB1	PB2
α helix	33%	23%
β pleated sheet	26%	30%
β reverse turn	23%	20%
Undefined structure	18%	27%

The secondary structure analysis was done using the Chou-Fasman Probabilistic method [19 a] as shown in Fig. 2 and Fig. 3. The content of secondary structural elements in PB1 and PB2 proteins were calculated as follows: (number of amino acids in a secondary structural element/total number of amino acids) × 100.

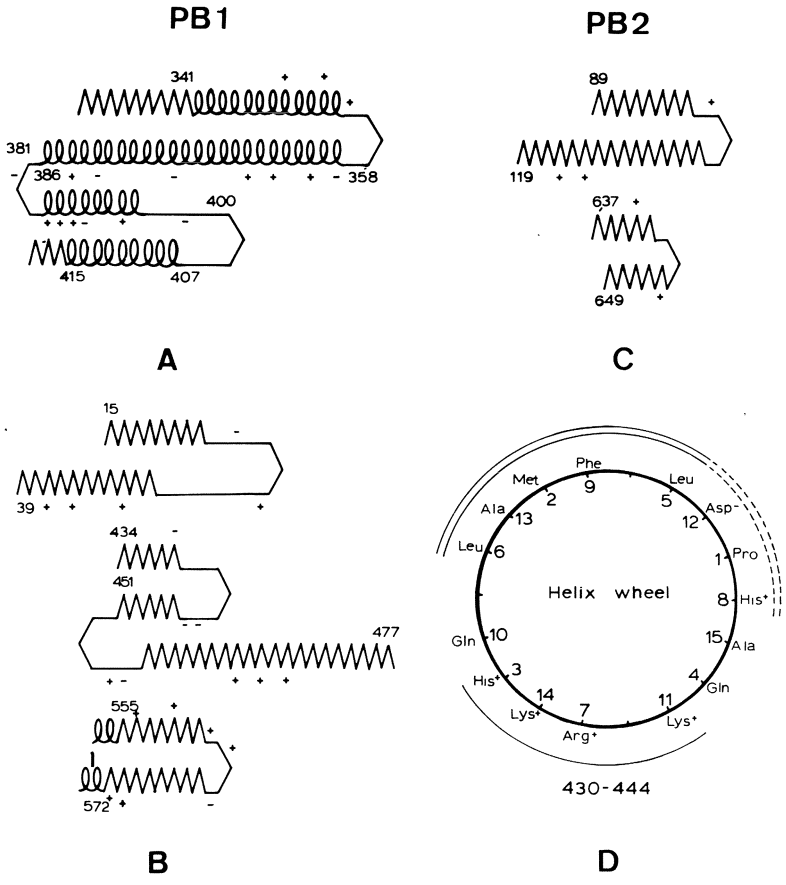


Fig. 4. Important secondary structural features of PB1 and PB2 proteins. *A* The four helical super-secondary structure in the PB1 protein. The amino acid residues between positions 341 and 415 [58] were predicted to form four antiparallel α-helices interrupted by β reverse turns. ⌇⌇⌇⌇⌇⌇⌇⌇⌇⌇ represents α-helices. β-turns are represented by chain reversals. *B* "Polypeptide double helices" (antiparallel β sheets) in the PB1 protein. Four antiparallel β sheets were predicted from the amino acid sequence [58] of the PB1 protein (Fig. 2). They are (a) amino acid residues between positions 15 and 39, (b) amino acid residues between 434 and 451, (c) amino acid residues between 447 and 477, (d) amino acid residues between 555 and 572. ⩙⩙⩙⩙⩙⩙⩙ represents amino acid residues which are involved in β-pleated sheets. β turns are represented by chain reversals. *C* "Polypeptide double helices" (antiparallel β sheets) in the PB2 protein. Two antiparallel β sheets were predicted from the amino acid sequence [36] of the PB2 protein (Fig. 3). They are (a) amino acid residues between positions 89 and 119 and (b) between 637 and 649. *D* "Helix Wheel" plot of one of the amphiphilic helices (amino acid residues 430—444) of PB2 protein. Numbers 1 to 15 represent the positions of amino acid residues from 430 to 444. The amino acid residues were plotted on a "helix wheel" according to the procedure of Schiffer and Edmundson [56 a] using 3.6 amino acid residues per turn. ————, the region of hydrophilicity and ═══, the region of hydrophobicity. =====, the region of hydrophobicity generated via a salt linkage between asp^- and his^+

protein (amino acid residues 26 to 33 and 430 to 444, Fig. 4 D) which possess larger hydrophobic moments than those of typical globular protein α-helices [58]. One of these α-helices when plotted on a "helix wheel" show that one face is moderately hydrophilic (His^+, Lys^+, Arg^+ and Lys^+, Fig. 4 D), while the other is moderately hydrophobic (Leu, Ala, Met, Phe and Leu, Fig. 4 D). Such α-helices having both hydrophilic and hydrophobic faces are called amphiphilic helices. The helical hydrophobic moment and mean hydrophobicity values of these helices are similar to the "surface-seeking helices" of mellitin, δ-hemolysin and diphtheria toxin polypeptides which are surface-active and lytic [71]. These amphiphilic helices are likely to occur on the surface of the protein and cause reversible association between proteins and surfaces of lipid bilayers.

V. Structure of DI RNAs

A. Classes of DI RNAs

Based on their primary structures, DI RNAs can be classified into four major classes (Fig. 5, ref. [40]): (i) 5′ DI RNA: DI RNAs belonging to this class retain the 5′ terminus of the genomic RNA but not the 3′ genomic terminus. The 3′ terminus of DI RNA is created by copying the 5′ terminus of DI RNA. Thus, the 5′ and 3′ termini of DI RNA become complementary to each other and can form a double-stranded stem of varying length. Diversity in this class of DI RNAs is created from the varying amount of deletion and the extent of double strand stem formed by copying the 5′ genomic terminus. These DI RNAs, therefore, lack the genomic transcriptase recognition site and the leader sequence at its 3′ terminus and, therefore, are unable to act as templates for transcription of poly(A)-containing capped messages, although a small RNA fragment is often transcribed [30]. The new polymerase binding site generated at its 3′ terminus is presumed to have a higher affinity for viral polymerase (replicase) than the 3′ genomic terminus and therefore, interferes with the replication of standard viral RNA by competing more effectively for the limited number of polymerase molecules. The majority of the DI RNAs from nonsegmented negative strand viruses belong to this class [30]. Although most of these DI RNAs have only one deletion point, the possibility of multiple deletions in some DI RNAs cannot be ruled out. (ii) 3′ DI RNA: These DI RNAs will

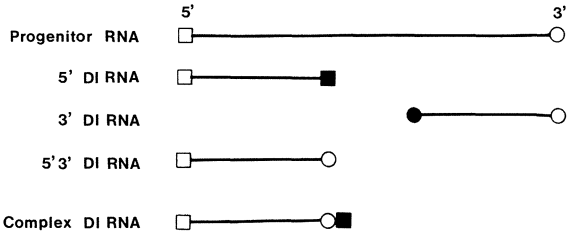


Fig. 5. Classes of DI RNAs. □, ○ represent genomic 5′ and 3′ termini respectively and ■, ● represent sequences complementary to the genomic 5′ and 3′ terminal sequences

be the counterpart of the 5′ DI RNA, *i.e.*, they should contain the 3′ terminus but lack the 5′ terminus of the genomic RNA. The 5′ terminus of this DI RNA will be created by copying the 3′ terminus of DI RNA. However, no such DI RNA has been reported to date, suggesting that the 5′ terminus of the genomic RNA is indispensable for virus replication and morphogenesis. In addition to polymerase binding activity, the sequence at the 5′ terminus may have other properties, such as the nucleation signal for nucleoprotein assembly and formation of virion, etc. This class of DI RNA, if found, is expected to transcribe poly(A)-containing capped messages in a manner analogous to that of the genomic RNA. (iii) 5′—3′ DI RNA: These DI RNAs contain internal deletion(s) but retain both the 5′ and the 3′ genomic termini and are expected to be transcribed into mRNA molecules. The majority if not all of the influenza DI RNAs [24, 51, 59] and some of the Sendai and VSV (vesicular stomatitis virus) DI RNAs belong to this class [4, 53]. (iv) Complex DI RNAs: Any DI RNAs that do not belong to one of the three classes mentioned above will belong to this group. Here, extensive changes occur in the DI RNA with the creation of new sequences and/or new termini. VSV DI LT2 [38], 18S DI RNA of Semliki Forest virus [41] and mosaic RNA of influenza [43] are examples of this class. The transcriptional properties of DI RNAs belonging to this class may vary, depending on a number of factors such as the nature of the 3′ terminus, the presence of transcription regulatory signals, transcription terminators, poly(A) addition sites, etc.

B. Complete Nucleotide Sequences of DI RNAs

Recently a number of DI RNAs of PB1 and PB2 origin have been completely sequenced, providing for the first time information about the specific relationship of the DI RNA to the progenitor gene at the nucleotide level. However, caution must be exercised in interpreting sequences obtained solely by DNA cloning, because of the possibility of sequence rearrangement during cloning [26].

Fig. 6 shows the complete sequence of two DI RNAs (L2b and L3) of PB1 origin. L2b and L3 contain 683 and 441 nucleotides, respectively, as expected from their migration in RNA gels (Fig. 1). In the plus sense, L2b contains 270 nucleotides from the 5′ terminus and 413 nucleotides from the 3′ terminus of the PB1 gene with a deletion of 1658 nucleotides. L3, on the other hand, contains 197 nucleotides from the 5′ terminus and 244 nucleotides from the 3′ terminus of the PB1 gene with an internal deletion of 1900 nucleotides. The sequence of L3 is a subset of L2b sequence as expected [51]. Except for the deletion, no other rearrangement or base mutation was observed among L2b, L3 and PB1 sequences.

The complete nucleotide sequences of four DI RNAs of PB2 origin have been determined by DNA cloning (Fig. 7, ref. [59, 70]). Clone B is 444 nucleotides long and contains 190 and 254 nucleotides from the 5′ and 3′ termini of the PB2 gene, respectivelly [70]. Clone C is 480 nucleotides long, containing 244 nucleotides from the 5′ terminus and 236 from the 3′ terminus of the PB2 gene plus strand [70]. Clones B and C were obtained from PR/8/34 virus and a few base mismatches were found in these sequences when compared to the progenitor PR/8 PB2 sequence;

Fig. 6. Nucleotide sequence of L2b and L3 DI RNAs of PB1 origin [51]. ————, regions that are absent in L2b and L3 DI RNAs; N - N - N - N represent a sequence of 1658 nucleotides in PB1 gene [58]

Fig. 7. Nucleotide sequence of DI RNAs of PB2 origin. □, the mismatches with the progenitor genes; ————, the common nucleotide sequence (GGA) at the deletion point [59, 70]; ————, regions that are absent in the DI RNAs; N - N - N - N, a sequence of 1626 nucleotides in PB2 gene [36]. L2a-7, L2a-17 are of WSN virus origin [59] whereas C and A, B were obtained from PR/8 virus [70]

A→C and A→U at nucleotides 94 and 369 respectively in clone B and A→C, A→C and U→A at nucleotides 94, 142, and 369 respectively in clone C. Two other DI RNA clones contain 659 (L2a-7) and 611 (L2a-17) nucleotides [59], which corresponds to the expected size of L2a RNAs of the WSN PB2 gene (Fig. 1). Both L2a-17 and L2a-7 contain 272 nucleotides from the 5′ terminus of the WSN PB2 gene and a deletion of 1682 nucleotides (nucleotide 273 to 1954). However, L2a-17 contains an additional deletion of 48 nucleotides of the P3 sequence (nucleotide 2032 to 2079). L2a-7 contains two mutations at nucleotide position 103 (G→C) and 423 (G→A) whereas L2a-17 contains a single mutation (G→A) at nucleotide 497.

These analyses show that influenza virus DI RNAs contain internal deletion(s) of various lengths. Furthermore, although the majority of them contain a single internal deletion, some contain multiple internal deletions. Except for the deletion(s) and a few base mismatches, extensive rearrangement of genomic sequences often found in other DI RNAs [41] has not been observed in influenza virus DI RNAs. However, more DI RNAs must be sequenced before definitive generalizations about the structure of influenza virus DI RNAs can be made.

VI. Generation of DI RNAs from the progenitor RNAs

Small RNA segments resembling DI RNAs are present in most if not all influenza virus preparations and virus stocks used in the laboratory. They are found in virus yields obtained from tissue culture plaques as well as in clonal stocks, albeit at a low level [35]. Although DI particles and DI RNAs are generally present in many influenza virus preparations, the mechanism of the generation and evolution of DI RNAs from their progenitor RNAs is a complex phenomenon and as yet, it is poorly defined for a number of reasons: (i) Until recently, the complete sequences of DI RNAs and their progenitor RNAs were not available and therefore, the precise sequence relationship of DI RNAs to the progenitor RNAs could not be determined. Furthermore, the number of sequences presently available is not sufficient for proposing a general model for the generation of DI RNAs. (ii) Moreover, DI RNAs that have been sequenced or studied extensively usually represent the predominant species of DI RNAs. Obviously, these DI RNAs have survived through selection pressure and have undergone many cycles of amplification. Therefore, they may not represent the first product in the generation of DI RNA from the progenitor gene. Again, they may not be the final DI RNA either, since it is known that the nature of the DI RNA changes with passage. Although it is not clear whether newly emerging DI RNAs arise from larger DI RNAs or from progenitor genes, neither of these possibilities can be excluded at present. Indeed, many of these DI RNAs may represent intermediate species in the evolution of DI RNAs. If multiple mechanisms are functioning in the different phases of generation and evolution of DI RNAs, it may be difficult to sort these out and propose a general model encompassing the entire process of genesis and evolution of DI RNAs. To determine the precise pathway and locate the intermediates in the generation of DI

Table 5. *Template sequences at the deletion sites in Influenza virus DI RNAs*

			Reinitiation site	Detachment site	
cRNA					
	L3	5′	ACAACAAA ↓CACCGAAA	ACCAAAAG ↓UGCUGCAA	3′
	L2b	5′	GUGGUUAU↓GCCCAAAC	AACCAUAA ↓AGACAUUG	3′
	L2a-7	5′	AAAAUGAA ↓UGACGCCG	GAGGGGAU ↓CAGGAAUG	3′
	L2a-17	5′	AAAAUGAA ↓UGACGCCG	GAGGGGAU ↓CAGGAAUG	3′
	L2a-17	5′	CAGUUCUC↓GGAAAGGA	GCACAGCU ↓GGAGUUGA	3′
	A & B	5′	AAUGAAAU ↓AUCCAAUU	GGAGUGGA ↓AUUCCUCA	3′
	C	5′	AAAUGAGC ↓AAGGACAA	CUGAGGGG ↓AUUCCUCA	3′
vRNA					
	L3	5′	UUGCAGCA↓CUUUUGGU	UUUCGGUG↓UUUGUUGU	3′
	L2b	5′	CAAUGUCU↓UUAUGGUU	GUUUGGGC↓AUAACCAC	3′
	L2a-7	5′	CAUUCCUG↓AUCCCCUC	CGGCGUCA↓UUCAUUUU	3′
	L2a-17	5′	CAUUCCUG↓AUCCCCUC	CGGCGUCA↓UUCAUUUU	3′
	L2a-17	5′	UCAACUCC↓AGCUGUGC	UCCUUUCC↓GAGAACUG	3′
	A & B	5′	CAGCGGAC↓UCCACUCC	AAUUGGAU ↓AUUUCAUU	3′
	C	5′	UGAGGAAU↓CCCCUCAG	UUGUCCUU↓GCUCAUUU	3′
Consensus splicing sequence			AG↓GURAG	Y-YYY-CAG ↓	

The template sequences [51, 59, 70] are tentatively designated as polymerase detachment sites and reinitiation sites depending on the synthesis of − or + strand. Template sequences are read from 3′ to 5′ direction. Arrows indicate the point of deletion. "R" and "Y" represent purine and pyrimidine base respectively, in the consensus splicing site [57 a].

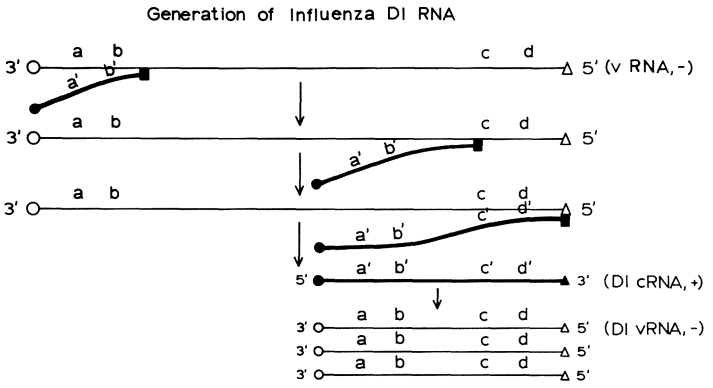


Fig. 8. Generation of internal deletion of a progenitor gene (or precursor DI RNA) via aberrant replication process. Polymerase molecule with attached nascent RNA strand may detach from the vRNA (negative strand) template, reattach down stream and, thereby, generate a shortened plus strand. These plus strand molecules could then serve as the template for the synthesis of deleted minus strands (DI RNA). Although not shown here, deletion can also take place during the synthesis of minus strand from a complete plus strand template. Again, once deleted minus strand is generated it can be amplified via the synthesis of plus strand template. ■, the polymerase molecule with the attached nascent RNA; *a*, *b* and *c*, *d* represent the 3′ and 5′ terminal sequences of the vRNA template

RNAs an *in vitro* system of viral and DI RNA replication would be necessary but no such system of replication is yet available. (iii) Additionally, information about the secondary structure of large RNAs, particularly as they exist in ribonucleoprotein complexes, is lacking. Such secondary structures, in addition to the sequence, may provide important clues regarding the regulation of replicational events, and therefore, the generation of DI RNAs. (iv) A further difficulty stems from the observations that influenza DI RNAs originate predominantly if not exclusively from the polymerase genes and sequence analyses have not revealed any unique structural features of the polymerase genes to account for this idiosyncrasy. Furthermore, it has not been resolved whether the paucity of non-polymerase DI RNAs is due to inability of these DI RNAs to be generated or to be able to survive and replicate efficiently during the amplification steps.

With these limitations in mind, our current knowledge about the generation and evolution of influenza virus DI RNAs can be summarized as follows. Sequence studies suggest that splicing is probably not involved in the generation of DI RNAs (Table 5). Sequences at the flanking regions of progenitor RNAs giving rise to different DI RNAs do not resemble the consensus cellular RNA splicing sequences [27, 51, 59]. Although splicing has been proposed to be involved in the generation of some of the messenger RNAs of M and NS genes, these genes have not yet been shown to produce DI RNAs. Similarly, sequences resembling both acceptor and donor splicing sequences are present in polymerase genes [36], but the junction points in DI RNAs do not involve these sequences (Table 5). Therefore, most likely, aberrant replication events during the synthesis of either the positive or the negative strand RNA involve skipping over a part of the RNA template (Fig. 8). However, sequence studies do not reveal a consensus sequence(s) for either the detach-

ment or the reattachment of polymerase during the genesis of DI RNAs (Table 5). The possible involvement of a uracil-rich region [27] in the template as a general or the only signal for polymerase detachment is untenable for a number of reasons: (i) Such sequences are not present at the junction points of all DI RNAs (plus or minus strand) (Table 5). (ii) Uracil-rich sequences are also present in M, NS, NA, HA and NP segments which do not give rise to the major DI RNAs (Table 1). Since influenza virus RNAs in general are uracil-rich [45], the presence of uracil clusters in sequences at the proximity of some junction points may be fortuitous. (iii) Furthermore, the pausing of RNA polymerase(s) in the uracil-rich sequences often causes slippage and addition of more adenine residues, as seen during the polyadenylation of viral messenger RNAs [3, 15]. However, no such additional adenine or other residues are found at the junction points in DI RNAs.

The majority if not all of the DI RNAs are of monogenic and not of polygenic origin, suggesting that the polymerase, with the nascent chain attached, may not completely detach from the template but possibly roll on to a new site of the template brought into juxtaposition by the formation of transient secondary structure(s). Such transient secondary structures may be formed during replication because of the dynamic nature of RNA-protein interactions. Although recombination among RNA molecules has been recently demonstrated [38a], the involvement of such a process generating influenza virus DI RNAs is apparently rare. Possible generation of one DI RNA involving complex recombinational events between PB1 and PB2 genes has been proposed, based on sequences obtained by recombinant DNA cloning [43].

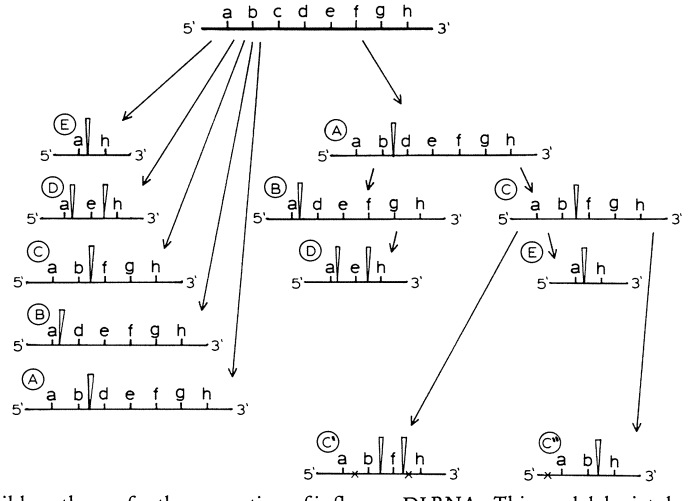


Fig. 9. Possible pathways for the generation of influenza DI RNAs. This model depicts how several DI RNAs *(A, B, C, D, E)* can be generated from a single progenitor gene using multiple pathways. In one case (left side) DI RNAs *(A, B, C, D, E)* containing varying internal deletions can be generated directly and independently from the same progenitor gene. Alternatively, the same DI RNAs can be generated through successive internal deletions of precursor DI RNAs (right side). Also, some DI RNA *(C)* can give rise to multiple DI RNAs *(C', C'')* through the accumulation of point mutations *(X)* as well as internal deletions (▽)

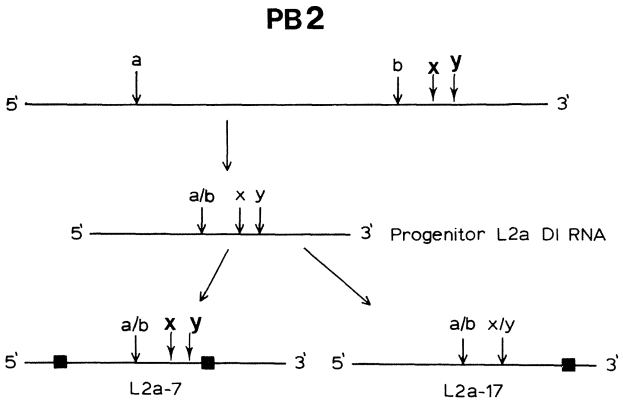


Fig. 10. Possible pathways for the generation and evolution of L2a-7 and L2a-17 from a progenitor DI RNA. *a*, *b*, *x* and *y* represent the points of deletion and ■ is the point of mutation [59]

Essentially there are two pathways in generating influenza virus DI RNAs from their progenitor genes. (i) All DI RNAs, large or small, are generated directly from their progenitor viral genes, *i.e.*, PB1, PB2, PA genes. (ii) Some of the DI RNAs, if not all of the major DI RNA species, do not originate directly from the progenitor gene(s) but from a precursor DI RNA. These two pathways are depicted in Fig. 9, showing how the same DI RNA can be generated either directly from the progenitor gene or from a precursor DI RNA. Clearly, some of the DI RNAs (the first generation product) must arise directly from the progenitor gene. The largest DI RNAs are good candidates for this pathway. However, even the largest DI RNAs studied to date do not possess all of the sequences present in some smaller DI RNAs (*e.g.*, L1 and $ts^{+}3$ or L1 and DIb of PB2 origin, ref. [48]). Therefore, either there must be an even larger precursor DI RNA or some of these small DI RNAs originate independently from the progenitor viral gene. There is a good possibility that some of the DI RNAs may arise from a precursor DI RNA and not directly from the progenitor viral gene, as shown in the case of L2a-7 and L2a-17 (Fig. 10). Since these two DI RNAs contain an identical deletion, it is unlikely that they arose independently from the progenitor PB2 gene. Rather, they may have come from a single precursor DI RNA which contained the common deletion and subsequently evolved into two DI RNA species—one with two base mismatches and the other with a separate base mismatch and an additional deletion (Fig. 10). This hypothesis suggests that DI RNAs undergo evolution after they are formed. In fact, the generation of progressively smaller DI RNAs after many passages has been reported [35]. Additionally, it was shown that some of the smaller DI RNAs contain sequences that are subsets of larger DI RNAs (*e.g.*, L2b and L3 of PB1 origin). These smaller DI RNAs may arise either by progressive alteration (including deletion) of the larger DI RNAs or directly from the progenitor viral gene (Fig. 11). Sequence data do not rule out either of these two possibilities. More likely, both pathways, independent generation directly from the progenitor viral gene(s), as well as progressive evolution may be operative in creating diverse influenza DI RNAs.

In summary, influenza DI RNAs are predominently, if not exclusively, generated from polymerase genes. A single polymerase gene can give rise to multiple DI RNAs of different sizes. One DI RNA usually originates from a single polymerase gene (*i.e.*, monogenic) which suggests that intersegmental recombination is rare. DI RNAs are generated from the progenitor gene(s) due to aberrant replication events. Furthermore DI RNAs once generated can undergo further evolution producing new DI RNAs.

VII. Transcription of Influenza Virus DI Viruses

Sequence studies showed that the structures of influenza virus DI RNAs are different from those of the 5′ DI RNAs observed with the nonsegmented negative strand viruses. Since influenza virus DI RNAs contained both the genomic 5′ and 3′ termini, they are likely to contain the genomic polymerase binding site at both ends. Based on this consideration, it was predicted that these DI RNAs should be transcribed into messenger RNA. Indeed, influenza DI viruses, unlike other negative strand DI viruses, transcribe complementary RNAs *in vitro* [17]. Transcription conditions were essentially identical to that of standard virus. Virion associated polymerase activity was also similar in both DI and standard virus preparations. As reported for standard virus [55, 56], transcription was dependent on the presence of specific external primers, such as ApG or capped RNAs. UpG, as expected, failed to serve as a primer for transcription. Since influenza virus DI particles, unlike nonsegmented RNA viruses, also contain standard RNA segments in addition to DI RNA segments, it was important to determine which of the RNA segments were transcribed. Fig. 12 shows that the sizes of the transcripts corresponds precisely to the sizes of the DI RNA segments. DI RNA transcripts also contain poly(A) (60–350 nucleotides long) at the 3′ end and ApG primer at the 5′ end [17]. Hybridization of DI RNA transcripts with the corresponding DI RNA

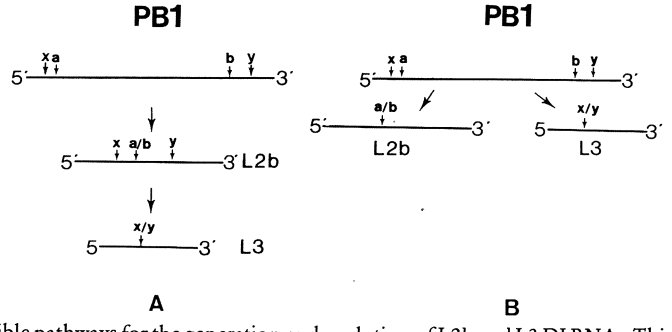


Fig. 11. Possible pathways for the generation and evolution of L2b and L3 DI RNAs. This figure shows how two DI RNAs (L2b and L3) which have been completely sequenced (Fig. 6) can arise via two independent pathways as proposed in Fig. 9. Since the entire sequence of L3 is present in L2b and since there is not a single base mismatch, L3 could have originated from the larger DI RNA, L2b *(A)*. Alternatively, both L2b and L3 could be generated independently from the progenitor gene PB1 *(B)*. *a*, *b*, *x* and *y* show the points of deletion

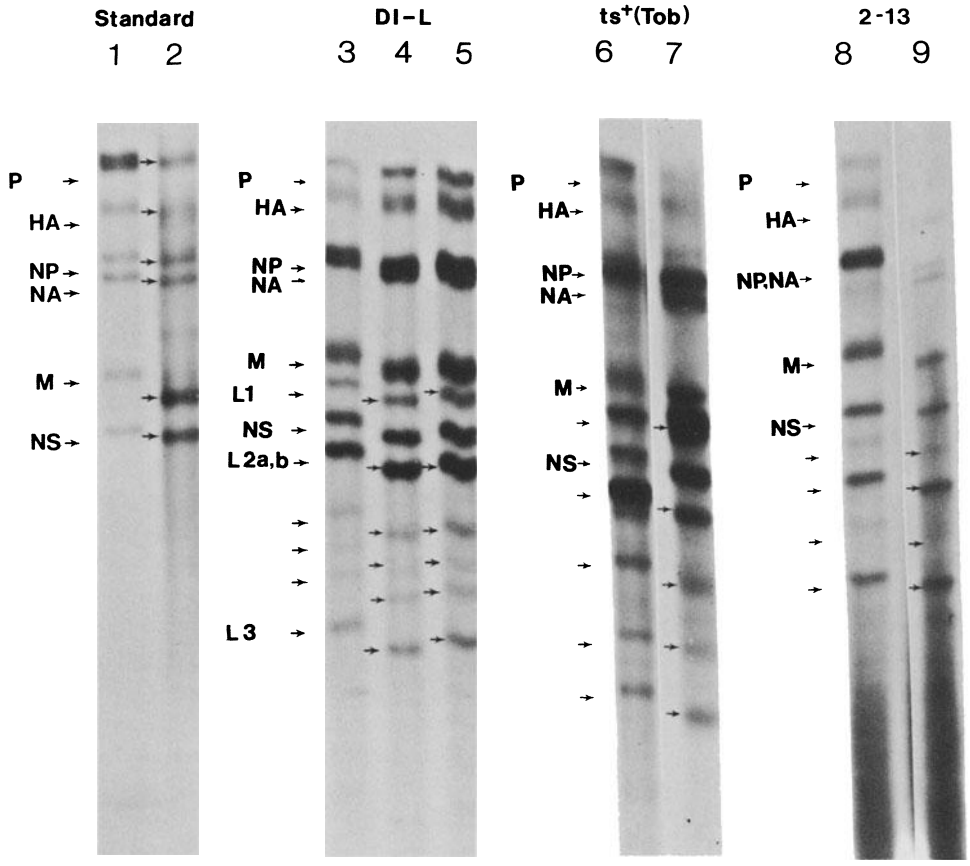


Fig. 12. PAGE analysis of transcripts of DI L, DI ts$^+$ (Tob) and DI 2-13 virus particles. Labeled poly (A)$^+$ transcripts were isolated, deadenylated and subjected to electrophoresis [17]. Lane *1*, standard genome RNA; Lane *2*, ApG primed standard genomic transcripts; Lane *3*, DI L RNA; Lane *4*, ApG primed DI L transcripts; Lane *5*, alfalfa mosaic virus RNA-4 primed DI L transcripts; Lane *6*, DI ts$^+$ (Tob) RNA; Lane *7*, ApG primed DI ts$^+$ (Tob) transcripts; Lane *8*, DI 2-13 RNA; Lane *9*, ApG primed DI 2-13 transcripts

segments and analysis of RNA duplexes demonstrated that DI RNA-specific transcripts are not incomplete polymerase gene products but represent faithful copies of the DI RNAs. These results clearly demonstrate that influenza DI RNAs are transcribed into poly(A)-containing complementary RNAs which have the characteristics of standard genomic transcripts.

Thus, from the transcriptional standpoint, DI RNAs can be classified into two groups: (i) those that cannot generate poly(A)-containing transcripts and (ii) others which can produce poly(A)-containing transcripts. The majority of influenza virus DI RNAs belongs to the second group, whereas most DI RNAs of nonsegmented negative strand viruses are of the 5′ type and belong to the first group. However, a recent report shows that 5′—3′ DI RNAs may not be uncommon among the nonsegmented viruses [4]; it remains to be seen if these DI RNAs can produce transcripts.

VIII. Interference by Influenza DI Viruses

The mechanism by which DI RNAs interfere with the replication of standard RNA is not fully understood. For the 5′ DI RNAs of nonsegmented negative strand viruses, it is postulated that since they lack the 3′ terminus of the genome but possess, at their 3′ terminus, alternative sequences with a higher affinity for the polymerase, they interfere with the replication of standard genome by competition for a limited number of polymerase molecules [54]. On the other hand, since influenza DI RNAs contain both the 5′ and the 3′ genomic termini and function as transcription templates, they are unlikely to possess altered polymerase binding sites at their termini. Therefore, the mechanism proposed for interference by the 5′ DI RNAs is not likely to be the major mode of interference for the 5′—3′ DI RNAs. Although a definitive pathway of interference by influenza DI viruses cannot be outlined at present, the data suggest a few possibilities which can be experimentally tested. First, the DI RNAs may interfere at the level of primary transcription. Indeed, such a mode of interference has been suggested for VSV DI LT, which is also of 5′—3′ type [57]. Secondly, since the DI RNA transcripts possess the characteristics of mRNAs, they may be translated, producing defective polymerase-like proteins. These proteins, if produced, may bind preferentially to the specific region(s) of genomic template(s) and inhibit transcription and/or replication of the genomic RNA(s). The majority of such defective proteins, will have the amino terminus of the corresponding polymerase proteins but a different carboxyl terminus because DI RNAs often do not maintain the reading frame of the progenitor gene. Finally, the transcription of DI RNAs may have no functional role in interference but other considerations such as structure or size of DI RNAs may be of significance in competition. Detailed *in vivo* studies of both the transcription of standard and DI RNA segments and the translation of genomic and DI transcripts will elucidate the mechanism of interference by 5′—3′ DI influenza viruses. Furthermore, DI particles mediated interference of influenza viruses may not only involve interference with the replication and/or transcription of viral RNA, but also interference at the level of assembly and maturation of the virus.

IX. Conclusion

Influenza DI particles produced at high multiplicity contain DI RNA molecules which are of the 5′—3′ type and arise from the internal deletions of one of the polymerase genes. They can be transcribed into poly(A)-containing complementary RNAs having the characteristics of standard viral mRNAs. The structure of influenza DI RNAs as well as their possible function (*e.g.*, transcription) appear to be different from that of the 5′ DI RNAs which represent the majority of nonsegmented negative strand DI RNAs. Although DNA cloning and sequencing have helped to elucidate the primary structure of a number of DI RNAs and their progenitor genes, the specific steps involved in the generation and evolution of influenza virus DI RNAs as well as the mechanism of interference remain largely unresolved and will be the subject of research for the next few years.

Influenza virus DI particles like DI particles of other viruses suppress the cytopathic effect of standard virus infection and aid in initiating persistently infected cultures *in vitro*. Since DI particles are ubiquitously present and interfere with standard viral replication, they may provide additional selection pressure on the evolution of influenza viruses in nature.

Acknowledgement

This work was supported by grants from the National Institute of Allergy and Infectious Diseases (RO1AI12749, RO1AI16348) and the National Science Foundation (PCM7823220). Authors gratefully acknowledge the assistance of Dr. Alan R. Davis in recombinant DNA cloning, DNA sequencing and many other phases of the work reported here.

References

[1] Ada, G. L., Perry, B. T.: Infectivity and nucleic acid content of influenza virus. Nature (Lond.) *175,* 209–210 (1955).
[2] Ada, G. L., Perry, B. T.: Influenza virus nucleic acid relationship between biological characteristics of the virus particle and properties of the nucleic acid. J. Gen. Microbiol. *14,* 623 (1956).
[3] Adman, R., Grossman, L.: Template properties of polyribonucleotides containing uracil or modified uracil in the RNA polymerase reaction. J. Mol. Biol. *23,* 417–439 (1967).
[4] Amesse, L. S., Pridgen, C. L., Kingsbury, D. W.: Sendai virus DI RNA species with conserved virus genome termini and extensive internal deletions. Virology *118,* 17–27 (1982).
[5] Bean, W. J., Simpson, R. W.: Transcription activity and genome conposition of defective influenza virus. J. Virol. *18,* 365–369 (1976).
[6] Bennink, J. R., Yewdell, J. W., Gerhard, W.: A viral polymerase involved in recognition of influenza virus-infected cells by a cytotoxic T-cell clone. Nature *296,* 75–76 (1982).
[7] Bishop, D. H. L., Huddleston, J. A., Brownlee, G. G.: The complete sequence of RNA segments of influenza A/NT/60/68 and its encoded P1 protein. Nucleic Acids Res. *10,* 1335–1343 (1982).
[8] Bishop, D. H. L., Jones, K. L., Huddleston, J. A., Brownlee, G. G.: Influenza A virus evolution: Complete sequences of influenza A/NT/60/68 RNA segment 3 and its predicted acidic P polypeptide compared with those of influenza A/PR/8/34. Virology *120,* 481–489 (1982).
[9] Brown, W. E., Stump, K. H., Kelley, W. S.: Escherichia coli DNA polymerase 1 sequence characterization and secondary structure prediction. J. Biol. Chem. *257,* 1965–1972 (1982).
[10] Camerini-Otero, R. D., Sollner-Webb, B., Felsenfeld, G.: The structure of the nucleosome: Evidence for an arginine rich histone kernel. In: Nucleic Acid–Protein Recognition (Vogel, H. J., ed.), 151–158. New York: Academic Press 1977.
[11] Carter, C. W., Kraut, J.: A proposed model for interaction of polypeptides with RNA. Proc. Natl. Acad. Sci. U.S.A. *71,* 283–287 (1974).
[12] Carter, M. J., Mahy, B. W. J.: Incomplete avian influenza virus contains a defective noninterfering component. Arch. Virol. *71,* 12–25 (1982).
[13] Carter, M. J., Mahy, B. W. J.: Synthesis of RNA segments 1–3 during generation of incomplete influenza A (Fowl plague) virus. Arch. Virol. *73,* 109–119 (1982).
[14] Carter, M. J., Mahy, B. W. J.: Incomplete avian influenza A virus displays anomalous interference. Arch. Virol. *74,* 71–76 (1982).
[15] Chamberlin, M., Berg, P.: Mechanism of RNA polymerase action. Characterization of DNA-dependent synthesis of polyadenylic acid. J. Mol. Biol. *8,* 708–726 (1964).
[16] Champness, J. N., Bloomer, A. C., Bricogne, G., Butler, P. J. G., Klug, A.: The structure of the protein disk of tobacco mosaic virus to 5 Å resolution. Nature *259,* 20–24 (1976).

[17] Chanda, P. K., Chambers, T. M., Nayak, D. P.: *In vitro* transcription of defective interfering particles of influenza virus produces poly(A)-containing complementary RNAs. J. Virol. *45,* 55–61 (1983).
[18] Choppin, P. W., Pons, M. W.: The RNAs of infective and incomplete influenza virions grown in MDBK and HeLa cells. Virology *42,* 603 (1970).
[19] Chou, P. Y., Adler, A. J., Fasman, G. D.: Conformational prediction and circular dichroism studies on the *lac* repressor. J. Mol. Biol. *96,* 29–45 (1975).
[19 a] Chou, P. Y., Fasman, G. D.: Empirical predictions of protein conformation. Annu. Rev. Biochem. *47,* 251–276 (1978).
[20] Crumpton, W. M., Dimmock, N. J., Minor, P. D., Avery, R. J.: The RNAs of defective-interfering influenza virus. Virology *90,* 370–373 (1978).
[21] Crumpton, W. M., Clewley, J. P., Dimmock, N. J., Avery, R. J.: Origin of subgenomic RNAs in defective-interfering influenza virus. FEMS Microbiol. Lett. *6,* 431–434 (1979).
[22] Crumpton, W. M., Avery, R. J., Dimmock, N. J.: Influence of the host cell on the genomic and subgenomic RNA content of defective-interfering influenza virus. J. Gen. Virol. *53,* 173–177 (1981).
[23] Davis, A. R., Nayak, D. P.: Sequence relationships among defective interfering influenza viral RNAs. Proc. Natl. Acad. Sci. U.S.A. *76,* 3092–3096 (1979).
[24] Davis, A. R., Hiti, A. L., Nayak, D. P.: Influenza defective interfering viral RNA is formed by internal deletion of genomic RNA. Proc. Natl. Acad. Sci. U.S.A. *77,* 215–219 (1980).
[24 a] Dayhoff, M. O.: Atlas of Protein Sequence and Structure, Vol. 5, Suppl. 3. National Biomedical Research Foundation, Georgetown University Medical Center, Washington, DC 20007, U.S.A., 363–373 (1978).
[25] De, B. K., Nayak, D. P.: Defective interfering influenza viruses and host cells: Establishment and maintenance of persistent influenza virus infection in MDBK and HeLa cells. J. Virol. *36,* 847–859 (1980).
[26] Fields, S., Winter, G.: Nucleotide sequence heterogeneity and sequence rearrangements in influenza virus cDNA. Gene *15,* 207–214 (1981).
[27] Fields, S., Winter, G.: Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structures of a small viral RNA segment. Cell *28,* 303–313 (1982).
[28] Fiers, W., Contreras, R., Duerinck, F., Haigeman, G., Iserentant, D., Merregaert, J., Minjou, W., Molemans, F., Raeymaekers, A., van den Berghe, A., Volckaert, G., Ysebaert, M.: Complete nucleotide sequence of bacterophage MS2 RNA: primary and secondary structure of the replicase gene. Nature *260,* 500–507 (1976).
[29] Garoff, H., Frischauf, A. M., Simons, K., Lehrach, H., Delius, H.: The capsid protein of Semliki forest virus has clusters of basic amino acids and prolines in its amino terminal region. Proc. Natl. Acad. Sci. U.S.A. *77,* 6376–6380 (1980).
[29 a] Hiti, A. L., Davis, A. R., Nayak, D. P.: Complete sequence analysis shows that the hemagglutinins of the H0 and H2 subtypes of human influenza virus are closely related. Virology *111,* 113–124 (1981).
[29 b] Hiti, A. L., Nayak, D. P.: Complete nucleotide sequence of the neuraminidase gene of human influenza virus A/WSN/33. J. Virol. *41,* 730–734 (1982).
[30] Holland, J. J., Kennedy, S. I. T., Semler, B. L., Jones, C. L., Roux, L., Grabau, E. A.: Defective interfering RNA viruses and the host cell response. In: Comprehensive Virology, Vol. 16 (Fraenkel-Conrat, H., Wagner, R. R., eds.), 137–192. New York: Plenum Press 1980.
[31] Huang, A. S., Baltimore, D.: Defective viral particles and viral disease processes. Nature (Lond.) *226,* 325 (1970).
[32] Huang, A. S., Baltimore, D.: Defective interfering animal viruses. In: Comprehensive Virology, Vol. 10 (Frankel-Conrat, H., Wagner, R. R., eds.), 73–116. New York: Plenum 1977.
[33] Irwin, M. J., Nyborg, J., Reid, B. R., Blow, D. M.: The crystal structure of tyrosyl-transfer RNA synthetase at 2.7 Å resolution. J. Mol. Biol. *105,* 577–586 (1976).
[34] Janda, J. M., Nayak, D. P.: Defective influenza viral ribonucleoproteins cause interference. J. Virol. *32,* 697–702 (1979).
[35] Janda, J. M., Davis, A. R., Nayak, D. P., De, B. K.: Diversity and generation of defective interfering influenza virus particles. Virology *95,* 48–58 (1979).

[36] Kaptein, J., Nayak, D. P.: Complete nucleotide sequence of the polymerase 3 (P3) gene of human influenza virus A/WSN/33. J. Virol. *42,* 55–63 (1982).
[37] Kavern, N., Kolomietz, L., Rudneva, I.: Incomplete influenza virus. Partial functional complementation as revealed by hemadsorbing cell count test. J. Virol. *34,* 506–511 (1980).
[38] Keene, J. D., Chien, I. M., Lazzarini, R. A.: Vesicular Stomatitis defective interfering particle contained a muted, internal leader RNA gene. Proc. Natl. Acad. Sci. U.S.A. *18,* 2090–2094 (1981).
[38 a] King, A. M. Q., McCahon, D., Slade, W. R., Newman, J. W. I.: Recombination in RNA. Cell *29,* 921–928 (1982).
[39] Kitamura, N., Semer, B. L., Rothberg, P. G., Larsen, G. R., Adler, C. J., Dorner, A. J., Emini, E. A., Hanecak, R., Lee, J. I., Van der Werf, S., Anderson, C. W., Wimmer, E.: Primary Structure, gene organization and polypeptide expression of polio-virus RNA. Nature *291,* 547–553 (1981).
[39 a] Lamb, R. A., Ching-Juh, Lai: Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of Influenza Virus. Cell *21,* 475–485 (1980).
[40] Lazzarini, R. A., Keene, J. D., Schubert, M.: The origins of Defective Interfering Particles of the negative strand RNA viruses. Cell *26,* 145–154 (1981).
[41] Lehtovaara, P., Soderlund, H., Keranen, S., Petterson, R. F., Kaarianen, L.: 18 S defective interfering RNA of Semliki forest virus contains a triplicated linear repeat. Proc. Natl. Acad. Sci. U.S.A. *78,* 5353–5357 (1981).
[42] Lenard, J., Compans, R. W.: Polypeptide composition of incomplete influenza virus grown in MDBK cells. Virology *65,* 418–426 (1975).
[43] Moss, B. A., Brownlee, G. G.: Sequence of DNA complementary to a small RNA segment of Influenza virus A/NT/60/68. Nucleic Acids Res. *9,* 1941–1947 (1981).
[44] Nakajima, K., Ueda, M., Sugiura, A.: Origin of small RNA in von Magnus particles of influenza virus. J. Virol. *29,* 1142–1148 (1979).
[45] Nayak, D. P., Baluda, M. A.: Isolation and partial characterization of nucleic acid of influenza virus. J. Virol. *1* (6), 1217–1223 (1967).
[46] Nayak, D. P.: Influenza viruses: structure, replication and defectiveness. Fed. Proc. *28,* 1858–1865 (1969).
[47] Nayak, D. P.: Defective virus RNA synthesis and production of incomplete influenza virus in chick embryo cells. J. Gen. Virol. *14,* 63 (1972).
[48] Nayak, D. P.: Defective interfering influenza viruses. Ann. Rev. Microbiol. *34,* 619–644 (1980).
[49] Nayak, D. P., Tobita, K., Janda, J. M., Davis, A. R., De, B. K.: Homologous interference mediated by defective interfering influenza virus derived from a temperature-sensitive mutant of influenza virus. J. Virol. *28,* 375–386 (1978).
[50] Nayak, D. P., Davis, A. R., Cortini, R.: Defective interfering influenza viruses: Complete sequence analysis of a DI RNA. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., ed.), 77–92. New York: Academic Press 1982.
[51] Nayak, D. P., Sivasubramanian, N., Davis, A. R., Cortini, R., Sung, J.: Complete sequence analyses show that two defective interfering influenza viral RNAs contain a single internal deletion of a polymerase gene. Proc. Natl. Acad. Sci. U.S.A. *79,* 2216–2220 (1982).
[52] Pasek, M., Goto, T., Gilbert, W., Zink, B., Schaller, H., Mackay, P., Leadbetter, G., Murray, K.: Hepatitis B virus genome and their expression in *Escherichia coli.* Nature (Lond.) *282,* 575–579 (1979).
[53] Perrault, J.: Origin and replication of Defective Interfering Particles. Curr. Topics in Microbiology and Immunology *93,* 152–207 (1981).
[54] Perrault, J., Semler, B. L., Leavitt, R. W., Holland, J. J.: Inverted complementary terminal sequences in defective interfering particle RNAs of vesicular somatitis virus and their possible role in autointerference. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 527–538. New York: Academic Press 1978.
[55] Plotch, S. J., Krug, R. M.: Influenza virus transcriptase: Synthesis *in vitro,* of large polyadenylic acid containing complementary RNA. J. Virol. *21,* 24–34 (1977).
[56] Plotch, S. J., Bouloy, M. Krug, R. M.: Transfer of 5′ terminal cap of globin mRNA during transcription *in vitro.* Proc. Natl. Acad. Sci. U.S.A. *76,* 1618–1622 (1979).

[56 a] Schiffer, M., Edmundson, A. B.: Use of helical wheels to represent the structures of proteins and to identify segments with helical potential. Biophys. J. *7,* 121–135 (1967).
[57] Schnitzlein, W. M., Reichmann, M. E.: The size and the cistronic origin of defective vesicular stomatitis virus particle RNAs in relation to homotypic and heterotypic interference. J. Mol. Biol. *101,* 307–325 (1976).
[57 a] Sharp, P. A.: Speculations on RNA splicing. Cell *23,* 643–646 (1981).
[58] Sivasubramanian, N., Nayak, D. P.: Sequence analysis of the polymerase 1 gene and the secondary structure prediction of polymerase 1 protein of human influenza virus A/WSN/33. J. Virol. *44,* 321–329 (1982).
[59] Sivasubramanian, N., Nayak, D. P.: Defective interfering influenza RNAs of polymerase 3 gene contain single as well as multiple internal deletions. Virology *124,* 232–237 (1983).
[60] Soeda, E., Arrand, J. R., Griffin, B. E.: Polyoma virus DNA: Complete nucleotide sequence of the gene which codes for polyoma virus capsid protein VP1 and overlaps the VP2-VP3 genes. J. Virol. *33,* 619–630 (1980).
[61] Ueda, M., Nakajima, K., Sugiura, A.: Extra RNAs of von Magnus particles of influenza virus cause reduction of particular polymerase genes. J. Virol. *34,* 1–8 (1980).
[62] Van Rompuy, L., Min Jou, W., Huylebroeck, D., Devos, R., Fiers, W.: Complete nucleotide sequence of the nucleoprotein gene from the human influenza strain A/PR/8/34 (HON 1). Eur. J. Biochem. *116,* 347–353 (1981).
[63] Von Heuverswyn, H., Van de Voorde, A., Fiers, W.: Nucleotide sequence of the Simian virus 40 Hind II & III restriction fragment J and the total amino acid sequence of the major structural protein VP1. Eur. J. Biochem. *91,* 415–430 (1978).
[64] Von Magnus, P.: Studies on interference in experimental influenza. I. Biological observations. Mineral. Geol. *24* (7), 1 (1947).
[65] Von Magnus, P.: Propagation of the PR-8 strain of influenza A virus in chick embryos. III. Properties of the incomplete virus produced in serial passages of undiluted virus. Acta. Pathol. Microbiol. Scand. *29,* 157 (1951).
[66] Von Magnus, P.: Incomplete forms of influenza virus. Adv. Virus Res. *2,* 59–78 (1954).
[67] Winter, G., Fields, S.: 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 (1980).
[68] Winter, G., Fields, S.: The structure of the gene encoding the nucleoprotein of human influenza virus A/PR/8/34. Virology *114,* 423–428 (1981).
[69] Winter, G., Fields, S.: Nucleotide sequence of human influenza A/PR/8/34 segment 2. Nucleic Acids Res. *10,* 2135–2143 (1982).
[70] Winter, G., Fields, S., Ratti, G.: The structure of the two subgenomic RNAs from human influenza virus A/PR/8/34. Nucleic Acid Res. *9,* 6907–6915 (1981).
[71] Eisenberg, D., Weiss, R. M., Terwilliger, T. C.: The helical hydrophobic moment: a measure of the amphiphilicity of a helix. Nature *299,* 371–374 (1982).

9

Influenza B and Influenza C Viruses

G. M. Air and *R. W. Compans*

I. Introduction

It has long been recognized that viruses classified as influenza can show a complete lack of cross-reactivity in their surface antigens. However, all viruses classified as influenza type A have cross-reacting internal components, the matrix and nucleocapsid proteins. Influenza B and C viruses are similar in morphology (Fig. 1) and composition, and in their replication processes, to influenza A viruses. However no antigenic cross reactivity among these viruses has been detected, and they are therefore classified as completely unrelated antigenic types. The morphology of influenza A and B virions is indistinguishable when examined by electron microscopy; on the other hand, influenza C virions exhibit some distinct structural features, which are described in detail in the text which follows.

Major differences are evident between the ecology and antigenic variability of influenza A and that of influenza B and C viruses. Influenza A viruses have been isolated from various animal species whereas influenza B and C viruses, except for a single report, have thus far been isolated only from man. No major antigenic shifts have been observed in the surface antigens of influenza B or C viruses, although antigenic drift occurs in the hemagglutinin of influenza B viruses.

This chapter describes what is known about the structure, replication, genetics and epidemiology of influenza B and C viruses, and their similarities to and differences from influenza A viruses.

II. Influenza B Virus

A. Influenza B RNA Species

In influenza B viruses there are 8 RNA segments (Ritchey *et al.*, 1976) which appear to code for at least 9 polypeptides analogous to those of the A viruses, and it has been determined which RNA segment codes for which viral protein(s) (Ueda *et al.*, 1978; Racaniello and Palese, 1979; Lamb and Choppin, 1979). The gene segments

of influenza B virus are all distinctly larger than the corresponding ones of influenza A viruses (Desselberger and Palese, 1978).

Sequence information on the genes and proteins of influenza B viruses is rapidly accumulating. The current knowledge is summarized in Table 1. Skehel and Hay (1978) reported sequences at the 5′ end of the vRNA of B/Hong Kong/8/73 and

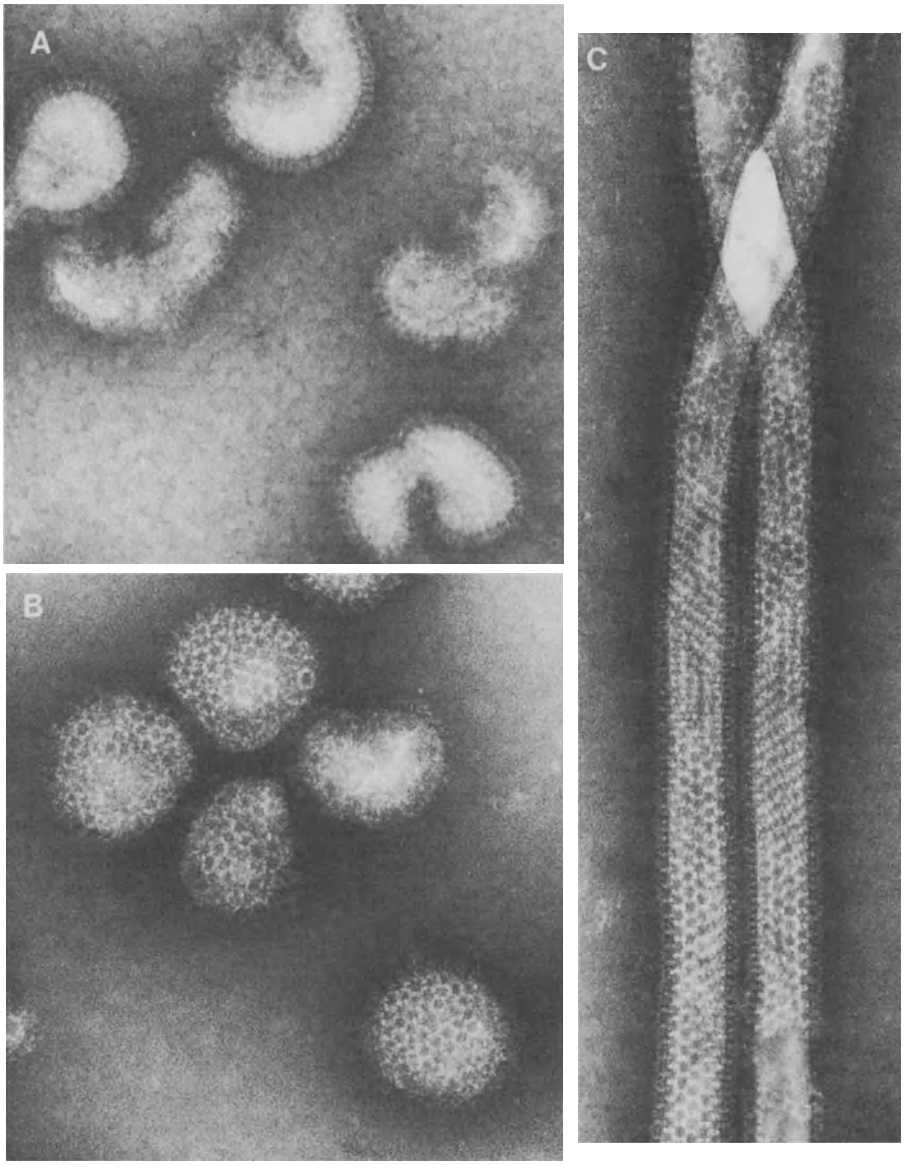


Fig. 1. Electron micrographs of influenza B and C virions. *A* Influenza B/HK/8/73 virions. *B* Influenza C/JHB/1/66 virions grown in embryonated eggs. *C* Filamentous influenza C/JHB/1/66 particles grown in chick embryo fibroblast cells. Negative staining with sodium phosphotungstate, pH 7.0. Magnification: A×150,000; B×150,000; C×165,000

Table 1. *Sizes of influenza B/Lee/40 genes and proteins*

Genes			Proteins			
Segment No.	No. of nucleotides (from gel electrophoresis)[a]	(from sequence)	Designation	MW (SDS gel electrophoresis)[a]	Carbohydrate	Amino acids (from nucleotide sequence)
1	2,800		P2 or P3	93,000 or 80,000		
2	2,800		P1	102,000		
3	2,700		P3 or P2	80,000 or 93,000		
4	2,100	1882[b]	HA	84,000	+	584[b]
			HA1		+	346
			HA2		+	223
5	2,000		NP	66,000		
6	1,700	1557[e]	NA	66,000	+	466[e]
7	1,150	1191[c]	M	25,000		248[c]
			M2			195[c]
8	1,000	1096[d]	NS1	40,000		281[d]
			NS2	11,500[d]		131[d]

[a] Racaniello and Palese (1979).
[b] Krystal *et al.* (1982).
[c] Briedis *et al.* (1982).
[d] Briedis and Lamb (1982).
[e] Shaw *et al.* (1982).

Desselberger *et al.* (1980) obtained 20 to 30 nucleotides of sequence from the 5′ and 3′ ends of the RNA segments of influenza B/Hong Kong/8/73. As in the case of influenza A viruses (Skehel and Hay, 1978), the RNA segments have common sequences at the ends, the B strain sequences being similar, but not identical to those of the A strains. At least in the hemagglutinin and neuraminidase genes, the larger size in B viruses is in part due to a longer non-coding sequence at the 5′ end of the vRNA (Krystal *et al.*, 1982; Shaw *et al.*, 1982).

From cloned DNA, complete sequences have been obtained for RNA segment 7 (Briedis *et al.*, 1982), segment 8 (Briedis and Lamb, 1982), segment 4 (Krystal *et al.*, 1982) and segment 6 (Shaw *et al.*, 1982) of B/Lee/40. The published terminal sequences of influenza B RNAs are shown in Tables 2 and 3, and suggest that there may be more heterogeneity than has been reported in the terminal sequences of influenza A viruses. In view of the relatively small antigenic variation in influenza B compared with influenza A, this is rather surprising.

B. Influenza B Proteins

Gel electrophoresis in the presence of SDS suggested that influenza B proteins have molecular weights which are different from those of their counterparts in influenza A viruses. The HA, NP and NA of the B/GL/1760/54 were found to be somewhat larger (82,000, 66,000, and 64,000 respectively), while the M protein was smaller (24,000) than the corresponding A/WSN/33 proteins (Choppin *et al.*, 1975; Compans *et al.*, 1970). Values very close to these were obtained in a later study of B/Lee/40 proteins (Racaniello and Palese, 1979).

1. Hemagglutinin

The influenza B HA protein undergoes proteolytic cleavage into HA1 and HA2 subunits, as do influenza A HA polypeptides. The cleavage can be mimicked by trypsin but apparently not by plasmin (Choppin *et al.*, 1975). The HA1 and HA2 polypeptides both contain glucosamine, but the HA2 from B/GL/1760/54 contains no fucose, whereas this sugar is present in HA2 of influenza A viruses (Choppin *et al.*, 1975), suggesting that the oligosaccharides linked to the influenza B HA2 subunit are of the simple, high mannose type.

Waterfield *et al.* (1979) sequenced the N-terminal regions of HA1 and HA2 of B/Lee/40, and showed that there is homology to the influenza A HA1 and HA2 sequences.

Krystal *et al.* (1982) have sequenced cloned cDNA copies of the HA gene of B/Lee/40. The complete HA sequence was derived from two clones, one extending from the synthetic oligonucleotide used as primer in cDNA synthesis to 1336 base pairs (bp). The second clone began at nucleotide 483 and extended to the 5′ end of the virion RNA, thus enabling the complete sequence to be determined. The total length is 1882 nucleotides. The first AUG begins at nucleotide 34, and when the sequence is translated from this initiation codon, a protein of 584 amino acids is predicted. The N-terminal sequence reported by Waterfield *et al.* (1979) begins at

amino acid 16. The sequence from 3 to 15 is non-polar and presumably is a signal peptide as in influenza A viruses (Elder *et al.*, 1979; Air, 1979).

The sequence of the N terminus of HA2 determined by Waterfield *et al.* (1979) begins after amino acid 346 of HA1, which is an arginine residue. Thus, proteolysis to cleave the precursor HA into HA1 and HA2 involves a trypsin-like activity, as in the case of influenza A viruses (reviewed by Ward, 1981) and as demonstrated by

Table 2. *3′-Terminal sequences of influenza B virus RNA segments*

Virus	Segment/Gene	Sequence	Reference
		10 20	
B/HK/8/73	1+2+3	UCGUCUUCGC U/G UCGUAAA-G	Desselberger *et al.* (1980)
B/HK/8/73	HA	UCGUCUUCGC UUCGUAAAAG	Desselberger *et al.* (1980)
B/Lee/40	HA	UCGUCUUCGC AACGUAAAAG	Krystal *et al.* (1982)
B/HK/8/73	NP	UCGUCUUCGC GUCGUAAAAG	Desselberger *et al.* (1980)
B/HK/8/73	NA	UCGUCUUCGC UUCGUAGAAG	Desselberger *et al.* (1980)
B/Lee/40	NA	UCGUCUUCGU CUCGUAUAAG	Shaw *et al.* (1982)
B/HK/8/73	M	UCGUCUUCGC UU	Desselberger *et al.* (1980)
B/Lee/40	7	UCGUCUUCGU GCGUGAAAGA	Briedis *et al.* (1982)
B/HK/8/73	NS	UCGUCUUCGC UUCCUAA U/C UA	Desselberger *et al.* (1980)
B/Lee/40	NS	U - - - - UUCGU CUCCUAAAUA	Air and Hackett (1980)
B/Lee/40	NS	(G)CGUCUUCGU CUCCUAAAUA	Briedis and Lamb (1982)

Table 3. *5′-Terminal sequences of influenza B RNA segments*

Virus	Segment	Sequence	Reference
		10 20	
B/HK/8/73	1+2+3*	AGUAGAAACA CGAGCAUUUU	Skehel and Hay (1978)
B/HK/8/73	1+2+3	AGUAGAAACA AGAG U/C AUUUU	Desselberger *et al.* (1980)
B/Lee/40	HA	AGUAGUAACA AGAGCAUUUU	Krystal *et al.* (1982)
B/HK/8/73	HA	AGUAGUAACA AGAGCAUUUU	Desselberger *et al.* (1980)
B/HK/8/73	4+5*	AGUAGAAACA AGAGCAUUUU	Skehel and Hay (1978)
B/HK/8/73	NP	AGUAGAAACA ACAGCAUUUU	Desselberger *et al.* (1980)
B/HK/8/73	6*	AGUAGAAACA AGAGCAUUUU	Skehel and Hay (1978)
B/HK/8/73	6	AGUAGUAACA AGAG U/C AUUUU	Desselberger *et al.* (1980)
B/Lee/40	6	AGUAGUAACA AGAGCAUUUU	Shaw *et al.* (1982)
B/HK/8/73	7*	AGUAGAAACA AG A/G GAAUUUU	Skehel and Hay (1978)
B/HK/8/73	7	AGUAGAAACA ACGCA C/A UUUU	Desselberger *et al.* (1980)
B/Lee/40	7	AGUAGAAACA ACGCACUUUU	Briedis *et al.* (1982)
B/HK/8/73	8*	AGUAGAAACA AGAGGAUUUU	Skehel and Hay (1978)
B/HK/8/73	8	AGUAGUAACA AGAGAAUUUU	Desselberger *et al.* (1980)
B/Lee/40	8	AGUAGUAACA AGAGGAUUUU	Briedis and Lamb (1982)

* Probable assignments

Choppin *et al.* (1975). It is not known whether further processing occurs at the C terminus of HA1.

Laver and Air (unpublished results) have analyzed tryptic peptides of HA1 and HA2 of B/Lee/40, which can be identified in the sequence predicted from the nucleotide sequence of Krystal *et al.* (1982). The sequences of HA1 and HA2 are shown in Fig. 2, with identified peptides marked as well as two differences in the sequences: Val was present in a peptide rather than Ala at position 90 of HA1, and Ser instead of Tyr at position 54 of HA2.

Potential glycosylation sites are present at residues 25, 59, 165, 232, 303, and 332 of HA1, and at 145, 171, 184, and 216 of HA2. In contrast, the influenza A hemagglutinins Mem/102/72 and Jap/305/57 have only one carbohydrate attached to HA2, and 6 and 4 respectively on HA1. In addition to the signal peptide, hydrophobic regions are seen at the N terminus and near the C terminus of HA2 (Krystal *et al.*, 1982). The C-terminal region of HA2 presumably is the transmembrane segment as in influenza A hemagglutinins (see Ward, 1981).

When the B/Lee/40 HA sequence is compared with that of A/PR/8/34 (Winter *et al.*, 1981), there is 24% and 39% conservation of amino acids in HA1 and HA2, respectively. When A/PR/8/34 is compared with A/Aichi/2/68 (Verhoeyen *et al.*, 1980) conservation is 35% and 53%; thus the B HA sequence is somewhat removed from, but still homologous to, the known influenza A HA sequences. Thirteen cysteine residues in B/Lee/40 and A/PR/8/34 can be aligned; the only extra cysteine residues are in the signal peptide and very close to the C terminus of HA2 of A/PR/8/34, neither of which is involved in the globular structure of the HA. Thus it is possible that the basic three-dimensional structure of B HA will be similar to that of A/Aichi/2/68 (Wilson *et al.*, 1981).

2. *Neuraminidase*

The only protein structure information on influenza B NA is an amino acid composition (Laver and Baker, 1972). However, an amino acid sequence can be predicted from the sequence of a cloned cDNA copy of viral and messenger RNA (Shaw *et al.*, 1982). The cloned DNA contains 10 nucleotides derived from cellular RNA by the "cap-transfer" process which is common to influenza A and B strains (Krug *et al.*, 1979).

The virus-specific portion of the B/Lee/40 NA clone sequenced by Shaw *et al.* (1982) is 1557 base pairs in length. The first AUG codon occurs at nucleotides 47 to 49, but the open reading frame for translation that begins at this codon terminates after nucleotide 346. This would yield a protein with a molecular weight of only 11,242, which is incompatible with the known size of the neuraminidase protein. The second AUG at nucleotides 54—56 is followed by an open reading frame extending to two consecutive termination codons at nucleotides 1452—1457, which suggests that this second AUG is the initiation site for the translation of the neuraminidase. Translation beginning at this codon would yield a protein consisting of 466 amino acids with a molecular weight of 51,721, which is the appropriate size for the neuraminidase. This protein would contain four potential glycosylation sites (*i.e.*, an asparagine followed by an unspecified amino acid,

```
                                                                  ⊢ HA1 ⟶
MET LYS ALA ILE ILE VAL LEU LEU MET VAL VAL THR SER ASN ALA ASP ARG ILE CYS THR  5
GLY ILE THR SER SER ASN SER PRO HIS VAL VAL LYS THR ALA THR GLN GLY GLU VAL ASN 25
VAL THR GLY VAL ILE PRO LEU THR THR THR PRO THR LYS SER HIS PHE ALA ASN LEU LYS 45
GLY THR GLN THR ARG GLY LYS LEU CYS PRO ASN CYS PHE ASN CYS THR ASP LEU ASP VAL 65
ALA LEU GLY ARG PRO LYS CYS MET GLY ASN THR PRO SER ALA LYS VAL SER ILE LEU HIS 85
GLU VAL LYS PRO ALA THR SER GLY CYS PHE PRO ILE MET HIS ASP ARG THR LYS ILE ARG 105
          VAL
GLN LEU PRO ASN LEU LEU ARG GLY TYR GLU ASN ILE ARG LEU SER THR SER ASN VAL ILE 125
ASN THR GLU THR ALA PRO ARG GLY SER TYR LYS VAL GLY THR SER GLY SER CYS PRO ASN 145
VAL ALA ASN GLY ASN GLY PHE PHE ASN THR MET ALA TRP VAL ILE PRO LYS ASP ASN ASN 165
LYS THR ALA ILE ASN PRO VAL THR VAL GLU VAL PRO TYR ILE CYS SER GLU GLY GLU ASP 185
GLN ILE THR VAL TRP GLY PHE HIS SER ASP ASP LYS THR GLN MET GLU ARG LEU TYR GLY 205
ASP SER ASN PRO GLN LYS PHE THR SER SER ALA ASN GLY VAL THR THR HIS TYR VAL SER 225
GLN ILE GLY GLY PHE PRO ASN GLN THR GLU ASP GLU GLY LEU LYS GLN SER GLY ARG ILE 245
VAL VAL ASP TYR MET VAL GLN LYS PRO GLY LYS THR GLY THR ILE VAL TYR GLN ARG GLY 265
ILE LEU LEU PRO GLN LYS VAL TRP CYS ALA SER GLY ARG SER LYS VAL ILE LYS GLY SER 285
LEU PRO LEU ILE GLY GLU ALA ASP CYS LEU HIS GLU LYS TYR GLY GLY LEU ASN LYS SER 305
LYS PRO TYR TYR THR GLY GLU HIS ALA LYS ALA ILE GLY ASN CYS PRO ILE TRP VAL LYS 325
THR PRO LEU LYS LEU ALA ASN GLY THR LYS TYR ARG PRO PRO ALA LYS LEU LEU LYS GLU 345
    ⊢ HA2 ⟶
ARG GLY PHE PHE GLY ALA ILE ALA GLY PHE LEU GLU GLY GLY TRP GLU GLY MET ILE ALA 19
GLY TRP HIS GLY TYR THR SER HIS GLY ALA HIS GLY VAL ALA VAL ALA ALA ASP LEU LYS 39
SER THR GLN GLU ALA ILE ASN LYS ILE THR LYS ASN LEU ASN TYR LEU SER GLU LEU GLU 59
                                                    SER
VAL LYS ASN LEU GLN ARG LEU SER GLY ALA MET ASN GLU LEU HIS ASP GLU ILE LEU GLU 79
LEU ASP GLU LYS VAL ASP ASP LEU ARG ALA ASP THR ILE SER SER GLN ILE GLU LEU ALA 99
VAL LEU LEU SER ASN GLU GLY ILE ILE ASN SER GLU ASP GLU HIS LEU LEU ALA LEU GLU 119
ARG LYS LEU LYS LYS MET LEU GLY PRO SER ALA VAL GLU ILE GLY ASN GLY CYS PHE GLU 139
THR LYS HIS LYS CYS ASN GLN THR CYS LEU ASP ARG ILE ALA ALA GLY THR PHE ASN ALA 159
GLY ASP PHE SER LEU PRO THR PHE ASP SER LEU ASN ILE THR ALA ALA SER LEU ASN ASP 179
ASP GLY LEU ASP ASN HIS THR ILE LEU LEU TYR TYR SER THR ALA ALA SER SER LEU ALA 199
VAL THR LEU MET ILE ALA ILE PHE ILE VAL TYR MET VAL SER ARG ASP ASN VAL SER CYS 219
SER ILE CYS LEU
```

Fig. 2. The predicted amino acid sequence of the hemagglutinin of B/Lee/40 deduced from the nucleotide sequence of cloned double-stranded DNA copies of the gene (Krystal *et al.*, 1982). Sequences at the N termini of HA1 and HA2 known from protein studies (Waterfield *et al.*, 1979) are indicated. Arrows underneath the sequence indicate tryptic peptides isolated from B/Lee/40 hemagglutinin with compositions matching the predicted sequence, except for the two amino acid differences shown (Laver and Air, unpublished results). Potential glycosylation sites (Asn-X-Thr/Ser) are shown in bold type. Only the Asn at position 332 of HA1 has been confirmed as glycosylated by the peptide studies

followed by serine or threonine), which are found at amino acids 56, 64, 144, and 284. The amino acid composition predicted from the nucleotide sequence is in close agreement with the amino acid analysis of purified influenza B/Lee/40 neuraminidase (Laver and Baker, 1972).

A striking feature of the amino acid sequences predicted from the nucleotide sequences of eight different subtypes of influenza A virus neuraminidase is that the first 6 amino acids at the N terminus are identical (Blok and Air, 1982 a, 1982 b). The amino acid sequence in this region of influenza B/Lee virus neuraminidase is different, although the third residue, proline, is the same in both the A and B proteins.

Influenza A neuraminidase has a region close to the N terminus which anchors the protein in the viral membrane (Fields *et al.*, 1981; Blok *et al.*, 1982), and similarly in the B/Lee/40 NA the only hydrophobic segment long enough to span the

Table 4. *Cysteine residues in neuraminidases of influenza A and B viruses*

B/Lee/40[a]	A/PR/8/34[b] and A/WSN/33[c]	A/Udorn/72[d] A/NT/60/68[e]
	14	21
		42
54	49	53
		78
87	77	92
122	109	124
127	114	129
	146	
		175
182	169	183
		193
229	216	230
231	218	232
236	223	237
251		
277	264	278
279	266	280
289	275	289
291	277	291
318	303	318
337	320	337
420	402	417
424	406	421
447	431	447

Residues are numbered from the initiating methionine residue which in some influenza A strains has been shown to be the N-terminal residue of the protein (Blok *et al.*, 1982).
Residue numbers on the same line are cysteines around which sequences show homology. Adapted from Shaw *et al.* (1982).

[a] Shaw *et al.* (1982).
[b] Fields *et al.* (1981).
[c] Hiti and Nayak (1982).
[d] Markoff and Lai (1982).
[e] Bentley and Brownlee (1982).

membrane extends from residues 4 to 34. Following this is a region which must correspond to the thin stalk seen in electron micrographs, and which in the B/Lee sequence (Shaw *et al.*, 1982) contains glycosylation sites and cysteine residues as was observed in the NA of various subtypes of influenza A (Blok and Air, 1982 a).

The NA sequences representing 8 subtypes of influenza A virus show no homology in the hydrophobic membrane-insertion sequence or the stalk (Blok and Air, 1982 a), but some homology becomes apparent about a cysteine which is at position 77 of PR/8/34 NA. Influenza B NA has no convincing homology with A/PR/8/34 or A/Udorn/72 NA until amino acid 116. Over the next 248 residues of B/Lee/40 NA, 87 amino acids are the same as those in A/Udorn/72 NA. Twelve of these are cysteine residues, suggesting that the 3-dimensional structures of the A and B neuraminidase may be basically the same, although A/Udorn/72 NA has two extra cysteine residues in this region (see Table 4).

The size of the influenza B virus neuraminidase protein, the amino-terminal hydrophobic region, and the homology between the A and B viruses, all suggest that translation of the B/Lee neuraminidase begins at the second AUG codon from the 5′ end of the mRNA (Shaw *et al.*, 1982). In the majority of eukaryotic mRNAs, including those of the other influenza virus genes sequenced, the first AUG codon is utilized as the initiation site for translation, but there are several cases in eukaryotic genes in which the first AUG codon is bypassed by the hypothesized "scanning mechanism" (Kozak, 1981). However, the influenza B NA gene is unusual in that there is a long open reading frame, which could code for 100 amino acids, following the first AUG codon. Synthesis of such a protein has not yet been reported (Shaw *et al.*, 1982).

3. Matrix Protein

The complete nucleotide sequence of a cloned full-length DNA copy of genome RNA segment 7 which codes for the matrix protein of influenza B/Lee/40 virus has been determined (Racaniello and Palese, 1979). The clone was derived from cDNA copies of viral RNA and messenger RNA (Briedis *et al.*, 1982).

RNA segment 7 of influenza B/Lee/40 virus contains 1191 nucleotides. After a 5′ noncoding region of 24 nucleotides, there is an open reading frame that extends from the first possible initiation codon for protein synthesis at nucleotides 25—27 to a termination codon at nucleotides 769—771. This region could code for a protein of 248 amino acids, which is compatible in size with the estimated molecular weight of the M protein. From nucleotide 513 there is a second open region, in the +2 reading frame, extending to a termination codon at nucleotides 1095—1097. A tract of five adenine residues occurs at nucleotides 1171—1175. This probably represents the polyadenylation site for the viral mRNA(s), as is seen in influenza A virus (Robertson *et al.*, 1981). Polyadenylation at this site would lead to an M_1 mRNA of approximately 1175 viral nucleotides.

The open region of the +2 reading frame could code for 195 amino acids which overlap those of the M_1 protein coding region by 86 amino acids. It is possible that the second coding region on RNA segment 7 of influenza B virus is used to code for a polypeptide analogous to influenza A virus M_2 (Briedis *et al.*, 1982).

Although this region could code for 195 amino acids, the first AUG codon does not occur until nucleotides 771—773. If protein synthesis were to initiate at this point, a protein of 105 amino acids would result. If, however, the influenza B M_2 mRNA contains an interrupted sequence and actually employs an initiation codon in a 5′ leader sequence, as is the case in influenza A virus (Lamb *et al.*, 1981), the body of the mRNA could code for any number of amino acids up to 195. No protein corresponding to an influenza B virus M_2 has yet been found either in infected cells or in hybrid-arrested translation experiments (Briedis *et al.*, 1981; Briedis and Lamb, 1982).

Influenza B virus M_1 protein contains 248 amino acids, including the initial methionine, while the A/Udorn/72 matrix protein contains 252 amino acids (Lamb and Lai, 1981). Sixty three of the 248 amino acids of the influenza B M_1 are common to the proteins of both viruses, but these are widely spaced and there is no sequence of more than three consecutive homologous amino acids. This observation may explain the lack of immunological cross-reactivity between the two proteins (Webster and Hinshaw, 1977) despite the 25% amino acid homology.

In influenza A M_1 protein, there is a long hydrophobic sequence between amino acids 115 and 151 (Winter and Fields, 1980; Lamb and Lai, 1981). In influenza B M_1 protein, a similar but shorter hydrophobic sequence exists between amino acids 122 and 151.

Three cysteine residues are conserved between the two strains in exactly the same positions in this region. Conservation of cysteine-containing internal hydrophobic regions within the M_1 proteins of influenza A and B viruses argues for their importance, which could relate to protein-protein or protein-lipid interactions for which sulfhydryl-bonded secondary structure is important.

After the end of the M_1 coding regions, the nucleotide homology between segment 7 of influenza A and B viruses falls to just the 1 in 4 expected to occur by chance, and no significant amino acid homology can be seen between the "M_2" sequences. The carboxyl terminus of the influenza B virus M_1 protein has a sequence of 9 amino acids after an AUG codon, and might be analogous to the potential peptide coded by the third influenza A segment 7 mRNA, which has only 9 amino acids, including the initiation methionine residue, which are identical to those at the carboxyl terminus of the M_1 protein.

Hybridization analysis (Scholtissek *et al.*, 1977) showed 51% base sequence homology between RNA segments 7 of influenza A and B viruses. The actual sequence homology is somewhat less than this, which demonstrates how the distribution of homologous nucleotides can affect hybridization results.

4. The Non-Structural Proteins

The eighth and smallest RNA segment of influenza B virus, like that of influenza A, codes for two non-structural proteins (Racaniello and Palese, 1979; Briedis *et al.*, 1981). The larger is designated NS_1, and the smaller, which is translated from a separate mRNA and shares no ^{35}S methionine-containing tryptic peptides with NS_1, is called NS_2 (Briedis *et al.*, 1981). The influenza A virus NS_1 and NS_2 polypeptides share nine amino acids at their N termini, and, after the leader sequence coding

for these amino acids, the NS_2 mRNA has an interrupted sequence of 473 nucleotides before its body, which is translated in the +1 reading frame (Lamb and Lai, 1980).

Double-stranded DNA clones derived from influenza B RNA segment 8 have been constructed by hybridization of full-length cDNA copies of RNA segment 8 and NS_1 mRNA. This DNA was cloned in plasmid pBR322 and sequenced (Briedis and Lamb, 1982). The NS_1 mRNA (1,080 viral nucleotides) contains nonviral nucleotides at its 5′ end and is capable of coding for a protein of 281 amino acids. Sequencing of NS_2 mRNA has shown that it contains an interrupted sequence of 655 nucleotides and is most likely synthesized by a splicing mechanism. The first $\approx$75 virus-specific nucleotides at the 5′ end of NS_2 mRNA are the same as are found at the 5′ end of NS_1 mRNA. This region contains the initiation codon for protein synthesis and coding information for 10 amino acids common to the two proteins. The $\approx$350 nucleotide body region of NS_2 mRNA can be translated in the +1 reading frame, and the sequence indicates that the NS_1 and NS_2 protein-coding regions overlap by 52 amino acids translated from different reading frames. Thus, between the influenza A and B viruses, the organization of the NS_1 and NS_2 mRNAs and the size of the NS_2 mRNA and protein are conserved, despite the larger size of the influenza B virus RNA segment 8, NS_1 mRNA, and NS_1 protein (Briedis and Lamb, 1982).

5. The P Proteins

Influenza B segments 1, 2, and 3 code for "P" proteins (Racaniello and Palese, 1979), which are presumably analogous to the three proteins coded by influenza A virus segments 1, 2, and 3 and which are concerned with transcription, replication, and cap transfer. The approximate sizes of the genes and proteins are included in Table 1, but so far no detailed information is available.

C. Replication of Influenza B Virus

Since the influenza B genome codes for proteins similar to those of influenza A viruses, it seems likely that the replication processes will be similar, and RNA-dependent RNA polymerase activity has been demonstrated in influenza B virions (Oxford, 1973). It is likely that, as in influenza A, transcription into mRNA results in a less than full length copy of the viral strand, with poly(A) at the 3′ end. Sequences of cloned cDNA show that at the 5′ end there are about 10 extra nucleotides and a cap structure derived from cellular mRNAs (Shaw *et al.*, 1982; Briedis *et al.*, 1982; Briedis and Lamb, 1982), as in influenza A (Krug *et al.*, 1979).

Apart from evidence that the nucleoprotein is synthesized early in infection and associates into virions with M protein synthesized late in infection (Choppin *et al.*, 1975), little information is available on the replication process.

D. Epidemiology

Virological surveillance of influenza epidemics over the past 25 years has shown that influenza A viruses were responsible for most of the epidemics where morbidity and mortality figures were high. However, in most winters since 1957, influenza B outbreaks have also occurred. Although influenza B epidemics have been reported to occur in 4 to 6 year cycles (Stuart-Harris and Schild, 1976), the disease may be more prevalent than the number of viruses isolated would indicate. Influenza B has most frequently been isolated from children, and since the disease does not usually result in admission to hospital, or death (Pereira, 1980), it may have been underestimated. However, over the 1979—1980 winter, there was a marked increase in the reported occurrence of influenza involving the B viruses. The disease affected adults as well as children, and was at times of greater severity than is normally associated with influenza B infection (Baine *et al.*, 1980).

The Northern winter of 1981—1982 was much lower in influenza outbreaks than the preceding two winters (Morbidity & Mortality Weekly Report, 1982), but 75% of all influenza isolates were influenza B. The disease predominantly affected school children, and although it was mild and did not result in detectable excess mortallity, there is cause for concern, because of the possible relationship between influenza B infection and Reye's syndrome (Corey *et al.*, 1977; Morbidity & Mortality Weekly Report, 1980 a).

Reye *et al.* (1963) described a clinical and pathological entity in Australian children which has since become recognized as a relatively common phenomenon. Increased awareness of the disease and new modes of treatment have reduced the mortality figures reported earlier of 85—100% to 35—50% (Schiff, 1976), but the disease is still a serious problem and its etiology remains to be determined. The primary symptoms are acute non-inflammatory encephalopathy, cerebral edema, and fatty degeneration of the liver. In most cases of Reye's syndrome there is an antecedent viral illness. The association with viruses is a strong one, although the pathological processes resemble the effects of toxins. Viruses could act merely as triggering mechanisms for exogenous or endogenous toxins, could exert a direct toxic effect, or could interact with toxins in a synergistic fashion. However, results of many epidemiological investigations and searches for potential toxins in patients and in their environment have so far been negative (Schiff, 1976), although aspirin has been strongly implicated (Morbidity and Mortality Weekly Report, 1980 b). The syndrome has been associated with multiple viruses, but the most frequent virus involved seems to be influenza B, and clusters of Reye's syndrome patients have appeared in relation to outbreaks of influenza B.

Studies on the epidemiology of Reye's syndrome in 1974 and 1977, when influenza B virus was prevalent, compared with the 1977—1978 period when influenza A viruses were dominant, have shown that in B-virus years there was approximately double the rate of occurrence of Reye's syndrome as in the A-virus year. Some evidence in 1977—1978 suggested that the influenza A virus was also associated with Reye's syndrome; however the relationship was not as marked as with influenza B viruses (Nelson *et al.*, 1979). Norman *et al.* (1968) obtained an influenza B virus from the liver of a Reye's syndrome patient. There have been

other, although rare, examples of influenza B isolation from tissues of patients suffering from this disease.

E. Antigenic Variation in Influenza B Viruses

Antigenic variation occurs in influenza B viruses, but is less marked than in influenza A viruses (Chakraverty, 1972; Schild *et al.*, 1973).

Recombinants between antigenically distinct influenza B viruses have been isolated in the laboratory with high frequency (Tobita and Kilbourne, 1974; Racaniello and Palese, 1979). However, so far, no major antigenic shifts involving influenza B viruses have been detected, and it is possible that influenza B recombinants may not have as dramatic epidemiological importance as recombinants between the A viruses. No influenza B virus has been reported from species other than man (Pereira, 1969), so it seems that influenza B viruses do not have alternate host reservoirs. Recombinants between influenza A and B strains have never been isolated.

Webster and Berton (1981) used monoclonal antibodies to the HA molecule of influenza B/Hong Kong/8/73 to investigate antigenic drift in these viruses more precisely. Antigenic drift occurred in each of the epitopes defined by 12 different monoclonal antibodies in influenza B viruses isolated from people since 1973. The sensitivity of the monoclonal antibodies permitted detection of antigenically distinguishable viruses from the same influenza B outbreak in Hannover in 1978 and 1979 and in Singapore in 1979.

These analyses were continued with monoclonal antibodies to the HA of B/Oregon/5/80 (Lu *et al.*, 1982). The majority of the 20 monoclonal antibodies did not inhibit influenza B viruses isolated before 1970, although some preparations did detect a single common determinant between B/Lee/40 and B/Ore/80. The majority of the monoclonal antibodies to B/Ore/80 reacted with B/Sing/222/79, but some differences were detectable. In studies using more than 50 influenza B strains isolated since 1975, it was shown that the 20 monoclonal antibodies displayed 13 different patterns of reactivity in H1 tests indicating that most of the antibodies were to different epitopes on the B/Ore/80 virus.

In 1982, epidemics of influenza B occurred in many countries of the world, being particularly severe in Japan and England. Analysis of recent isolates from these countries show that antigenic variants that react with only one or two of the 13 anti-B/Oregon/80 monoclonal antibodies were circulating (*e.g.*, B/Eng/19/82 and S/Shiga/75/82), suggesting that antigenic drift had occurred. On the other hand, other isolates from the same epidemic shared 10 of the 13 epitopes. Preliminary studies suggest that the majority of influenza B viruses isolated in Japan and England in 1982 react with only a limited number of the monoclonal antibodies whereas the majority of isolates from the USA react with many more of the monoclonal antibodies in the panel.

It is now apparent that many different antigenic variants of influenza B virus can be isolated during an epidemic year and that antigenically similar viruses are isolated in different countries. The majority of influenza B isolates were

distinguishable from each other; 19 isolates from Japan gave 14 patterns of reactivity with the panel of monoclonal antibodies. Nevertheless, viruses were isolated in Japan and England that had the same reactivity patterns with the monoclonal antibodies (Lu *et al.*, 1982).

One question is whether the multiple different strains co-circulate or whether during the course of an epidemic selection of antigenic variants occurs. Lu *et al.* (1982) report that during an epidemic in Memphis two distinct antigenic variants were isolated during the same week; B/Mem/1/82 shared 12 of 13 antigenic determinants with B/Ore/5/80 but Mem/3/82 shared only 4 of 13 determinants with B/Ore/5/80. The simplest explanation of these results is that multiple influenza B viruses co-circulate during an epidemic. A similarity between B/Mem/1/82 and B/Albany/1/77 also suggests that antigenic variants may circulate in humans for several years or alternatively that similar variants arise at different times from a parental strain.

So far, sequence changes involved in variation in influenza B viruses have not been characterized.

III. Influenza C Virus

Recent studies have indicated that influenza C viruses share many features of structure and replication with other orthomyxoviruses. However, some important differences have been recognized in the number of RNA genome segments and protein components, as well as in the type and distribution of biological activities in viral glycoproteins. These differences have provided an impetus for studies currently underway to characterize the genes and virus-coded polypeptides of influenza C viruses in more detail. Influenza C virus also differs in that no significant antigenic variation has been detected among virus isolates, in contrast to influenza A and B viruses. The reason for this difference is not apparent.

A. Virus Structure

1. Morphology of Virions

Influenza C virions exhibit many of the same structural features as influenza A and B virions, and both spherical particles and long filamentous forms are found in virus preparations (Fig. 1). A striking feature which distinguishes influenza C from most A and B virus preparations is the packing of surface glycoproteins in regular hexagonal arrays (Waterson *et al.*, 1963; Flewett and Apostolov, 1967; Compans *et al.*, 1977). At high magnification, a single morphological subunit is observed at each of the six vertices of the hexagonal arrays (Herrler *et al.*, 1981). It has been suggested that the hexagonal organization may involve direct lateral interactions between the glycoprotein molecules themselves, since the spikes are sometimes observed to maintain their arrangement in a network upon release from the viral membrane by limited proteolytic digestion or upon spontaneous disruption of the viral envelope

(Herrler *et al.*, 1981). Influenza C virus spikes have a length of 8–10 nm and a diameter of 4–5 nm. No convincing evidence for such a regular arrangement of spikes has been found on the surfaces of most influenza A viruses. As discussed below, only a single species of glycoprotein has been detected in influenza C virus, and this feature may be essential for formation of a regular surface lattice of glycoproteins. The presence of morphologically distinct hemagglutinin and neuraminidase spikes in influenza A and B viruses, which possess different elements of symmetry, may preclude the formation of a regular surface arrangement.

The ribonucleoprotein components of influenza C virions (Compans *et al.*, 1977) consist of short helical segments similar in dimensions to those obtained from detergent-disrupted influenza A virions (Pons *et al.*, 1969; Compans *et al.*, 1972). The strand-like structures measure about 15 nm in width, and range from 30 to 100 nm in length. By centrifugation in a rate zonal gradient, the ribonucleoproteins may be separated into multiple size classes, presumably containing RNA molecules of different sizes.

2. *Viral RNA Species*

The available information indicates that influenza C virions contain a segmented genome consisting of seven RNA species with molecular weights ranging from $0.94–0.23 \times 10^6$ (Petri *et al.*, 1979 c; Racaniello and Palese, 1980; Clerx *et al.*, 1983). Although no definitive information has been obtained to determine the coding assignment for each RNA segment, some features of the coding relationships between viral RNA species and polypeptides have been suggested from their molecular weight relationships and analogies to influenza A viruses, and are summarized in Table 5.

The 3′- and 5′-terminal sequences of influenza C viral RNA segments have been analyzed, and were observed to show partial homology with the corresponding sequences in influenza A and B viruses (Desselberger *et al.*, 1980). The 3′-terminal sequence as determined by pCp end labeling was found to be UCGU($\frac{U}{C}$)UUC-

Table 5. *Influenza C RNA and protein species*[a]

RNA segment	Approximate No. of nucleotides	Polypeptide Designation	Mol. Wt.
1	2350	P1	89,000
2	2350	P2	85,000
3	2150	P3	82,000
4	2000	gp88	64,000
5	1750	NP	60,000
6	1150	M	30,000
7	975	NS_1	25,000
		NS_2	14,000

[a] Data on sizes of RNA species are from Clerx *et al.* (1983). The coding relationships between RNA and protein species have not been established; both are listed in order of decreasing molecular weight.

GUCC for each of the RNA segments. In position 5, RNA segments 1, 2, 3, and 4 contained a U residue whereas segments 5, 6, and 7 contained a C residue; otherwise the terminal sequences of 12 residues were indistinguishable. After the twelfth nucleotide, some sequence divergence was evident. The 5′-terminal sequences were observed to be AGCAGUAGCAA for each of the RNA segments. The 3′-terminal sequences of influenza C RNA segments differ only in positions 4 and 5 from the corresponding 3′-terminal sequences of influenza A viral RNAs, and the 5′ sequences differ at positions 3, 6, and 8 between the two viral types. This similarity indicates a close evolutionary relationship of these viruses.

3. Viral Polypeptides

The internal structural polypeptides of influenza C virions closely resemble those reported for influenza A and B viruses; these include a nucleoprotein subunit (NP) and a membrane protein (M) as major polypeptide species as well as three minor polypeptides designated P1, P2, and P3 (Kendal, 1975; Compans *et al.*, 1977; Petri *et al.*, 1980). In contrast, however, to the two distinct glycoproteins of influenza A and B viruses which possess the hemagglutinin and neuraminidase activities, only a single type of glycoprotein has been detected in influenza C virions (Herrler *et al.*, 1979, 1981; Meier-Ewert *et al.*, 1980). This glycoprotein appears to possess receptor-binding activity, a penetration function, and an unidentified receptor-destroying enzyme activity (see below). It is designated gp88 based on its molecular weight (88,000) and because of the incomplete information concerning its function. The unglycosylated form of the glycoprotein was recently identified in tunicamycin-treated cells, and was found to have an estimated molecular weight of 64,000 (Nagele and Meier-Ewert, 1983). The gp88 glycoprotein migrates in the same region of an SDS polyacrylamide gel as the three P proteins; therefore, the latter were only identified following removal of the glycoproteins by treatment with the protease bromelain or by detergent extraction (Petri *et al.*, 1980). A viral RNA-dependent RNA polymerase activity has also been detected in influenza C virions, and by analogy with influenza A virus, is likely to be associated with the P polypeptides (Meier-Ewert *et al.*, 1981). Analysis of polypeptides synthesized in influenza C virus-infected cells also suggests that the genome of influenza C virus may code for two nonstructural proteins, designated NS_1 and NS_2, corresponding to the two products of the smallest RNA segment of influenza A virus (Petri *et al.*, 1980). A summary of the influenza C viral polypeptide species and their estimated molecular weights is presented in Table 5. Apart from the glycoprotein, the influenza C polypeptides have not been further characterized with regard to biochemical properties.

The influenza C glycoprotein, which is synthesized as a single polypeptide chain, is converted into two disulfide-linked subunits by proteolytic cleavage (Herrler *et al.*, 1979). Influenza C virus preparations from different host cells were observed to vary in the extent of glycoprotein cleavage. Virions grown in chick embryo fibroblasts (CEF) cells contained exclusively the uncleaved glycoprotein, whereas egg-grown virions possessed predominantly the cleaved form. Treatment of CEF-grown virions with trypsin resulted in cleavage of the glycoprotein *in vitro,*

and a concomitant increase up to 50-fold was observed in the specific infectivity of the virus preparation, indicating that the cleavage is essential for viral infectivity (Herrler *et al.*, 1979). These results are similar to previous observations with the hemagglutinin glycoprotein of influenza A virus (Klenk *et al.*, 1975; Lazarowitz and Choppin, 1975), in which cleavage of the HA polypeptide into two subunits by trypsin was observed to enhance the infectivity of virus preparations significantly. The activation of influenza C virus infectivity can also be accomplished by treatment with elastase, an enzyme of different specificity, suggesting that infectivity can be activated by cleavage at two nearby but different sites on the gp88 molecule (Sugawara *et al.*, 1981). In the same study, chymotrypsin and thermolysin were observed to be ineffective in cleavage of gp88 or activation of infectivity, and none of the enzymes tested were observed to affect the HA titer.

When analyzed under nonreducing conditions, the uncleaved glycoprotein differs significantly in its apparent molecular weight from the unreduced cleaved form; the two forms have apparent molecular weights of 105,000 and 85,000, respectively (Herrler *et al.*, 1981). The reason for this difference is not known. After reduction with mercaptoethanol, the uncleaved glycoprotein has an apparent molecular weight of 88,000, and the cleaved glycoprotein is separated into two subunits with apparent molecular weights of 65,000 and 30,000 (designated gp65 and gp30). The uncleaved glycoprotein appears to be more highly glycosylated, suggesting that a carbohydrate-containing portion may be removed during the proteolytic cleavage event (Herrler *et al.*, 1979). Also, a doublet of gp65 is frequently observed under reducing conditions, and the band with lower electrophoretic mobility also appears to be more highly glycosylated (Meier-Ewert *et al.*, 1981 a). However, no qualitative differences have been detected in the Pronase-derived glycopeptides of the two gp65 bands, nor has it been possible to detect differences in the tryptic peptides of these two bands, or of the uncleaved *vs.* the cleaved glycoproteins (Meier-Ewert *et al.*, 1981 a; Herrler *et al.*, 1981). Therefore, further studies are needed to define the nature of the proteolytic cleavage event and the molecular differences between the cleaved and uncleaved forms of the influenza C virus glycoproteins.

Analysis of N-terminal amino acid sequences indicated that the N termini of gp88 and gp65 are resistant to Edman degradation, suggesting the presence of blocked N termini (Herrler *et al.*, 1981). These results indicate that gp65 is derived from the N-terminal portion of the precursor glycoprotein. A partial sequence was determined for gp30, which was found to contain a preponderance of hydrophobic residues and was partially homologous with the corresponding sequence of the HA_2 subunit of influenza A viruses. In particular, the N-terminal tripeptide sequence of influenza C gp30 was found to be Ile-Phe-Gly, which corresponds exactly to a sequence observed in the HA_2 glycoprotein of A/Victoria/3/75 virus; however, the latter contains an additional N-terminal glycine residue (Min Jou *et al.*, 1980). A homologous sequence is also present at the N termini of the F_1 subunits of paramyxovirus F glycoproteins, and in these glycoproteins the terminal glycine residue is absent, thus resembling the terminal sequence found in influenza C gp30 (Gething *et al.*, 1978; Scheid *et al.*, 1978). These results, as well as the biological activity of the glycoprotein (see below), suggest that the influenza C glycoprotein

may exhibit structural features intermediate between those of influenza A glycoproteins and paramyxovirus glycoproteins.

4. Fusion and Hemolysis Activities

Activation of infectivity by proteolytic cleavage of gp88 was shown to be accompanied by an activation of hemolysis and fusion activities of influenza C virions (Ohuchi *et al.*, 1982). The hemolytic activity was readily detected with mouse erythrocytes whereas human and chicken erythrocytes were insensitive. Both the hemolysis and fusion activities required acidic pH, as has been observed for influenza A virus (Maeda and Ohnishi, 1980; Huang *et al.*, 1981; Lenard and Miller, 1981). The requirement for acid pH suggested that influenza C virus penetration involves uptake into lysosomes followed by fusion of the viral envelope and lysosomal membrane, as was originally proposed for Semliki Forest virus (Helenius *et al.*, 1980). These results suggest that cleavage of gp88 is essential for virus penetration, as is the case with the HA glycoprotein of influenza A virus and the F glycoprotein of paramyxoviruses.

5. Receptor Destroying Enzyme

Evidence for a difference between the receptor destroying enzyme of influenza C virus and other orthomyxoviruses was first obtained by Hirst (1950), who showed that influenza A and B viruses destroyed their own receptors on erythrocytes without affecting the receptor for influenza C virus. Conversely, influenza C virions elute from erythrocytes with destruction of their receptors, whereas receptors for other ortho- and paramyxoviruses are not removed (Kendal, 1975; Meier-Ewert *et al.*, 1978). Kendal (1975) has obtained further evidence that the receptor destroying enzyme of influenza C virions is not an α-neuraminidase, in that it did not liberate neuraminic acid from any of a series of known neuraminidase substrates. In addition, hemagglutination by influenza C virions was not inhibited by soluble neuraminic acid-containing glycoproteins that are active against influenza A viruses, whereas a rat serum inhibitor was observed to be uniquely active against influenza C virions (O'Callaghan *et al.*, 1977). The conclusion that influenza C virions lack a neuraminidase is also supported by the finding of sialic acid in the carbohydrate components of influenza C glycoproteins (Nakamura *et al.*, 1979). Three size classes of glycopeptides were observed upon digestion of influenza C virions or isolated glycoproteins with Pronase. The glycopeptides (designated G1, G2, and G3) were each detected in gp88, gp65, or gp30; however, the relative amounts of the three components varied among the glycoproteins. Both G1 and G2 were observed to be decreased in size by treatment with purified *Vibrio cholera* neuraminidase, and sialic acid was identified as the released product by thin layer chromatography. In contrast, influenza A viruses lack sialic acid (Klenk *et al.*, 1970; Schwarz *et al.*, 1977; Nakamura and Compans, 1978).

Thus far, attempts to identify the receptor and to elucidate the nature of the receptor destroying enzyme for influenza C virus have been unsuccessful. Evidence has been obtained that the receptor on chicken erythrocytes for influenza C virus is

inactivated by formaldehyde treatment, but not by periodate oxidation (Ohuchi *et al.*, 1978). Also, these authors observed that influenza C virus is able to agglutinate most human type B erythrocytes, but not type A erythrocytes. The nature of the viral receptor and the receptor destroying enzyme activity in influenza C virions remain as major unsolved problems in understanding the molecular biology of this virus.

B. Influenza C Virus Replication

Comparatively little information is available on the replication process of influenza C virus. The replication cycle has been studied in primary cultures of chicken kidney (CK) cells (Petri *et al.*, 1979 a) as well as several established lines (O'Callaghan *et al.*, 1977; Nerome and Ishida, 1978; Nerome *et al.*, 1979). Under single cycle conditions, hemagglutinin and infectious virus were first detected in CK cells around 8 hours post infection, and reached maximum titers at 24 hours post infection. This growth cycle is much slower than that of influenza A viruses in most host cell systems. Cellular polypeptide synthesis was not significantly inhibited by infection of CK cells with influenza C virus; however, the major viral polypeptides could be detected above the host cell background by comparison of infected and uninfected cells labeled in parallel. These polypeptides included the uncleaved glycoprotein (gp88) the nucleoprotein (NP) and the M protein, as well as two nonstructural proteins which have been designated NS_1 and NS_2 (mol. wts. 24,000 and 14,000). These nonstructural proteins are similar in size to those encoded by the smallest RNA segment of influenza A virus (Lamb and Choppin, 1979; Inglis *et al.*, 1979).

Like influenza A and B viruses, influenza C virus requires a host cell nuclear function for its replication process. The inhibitors actinomycin D and α-amanitin were observed to block the replication process of influenza C virus when added to infected cell cultures early in the growth cycle (Petri *et al.*, 1979 b). The actinomycin D-sensitive phase in replication was found to be more prolonged than that of influenza A virus, an observation consistant with the slower replication cycle observed for influenza C. Pretreatment of the host cell with UV light was also observed to block the replication of influenza C and influenza A viruses to the same degree (Petri *et al.*, 1979 b). A nuclear phase in influenza C virus replication has also been indicated by immunofluorescent studies showing that some influenza C antigens occur in nuclei of infected cells (Barclay *et al.*, 1971). The nuclear function in influenza A viruses appears to be related to the splicing of 5′ cap structures from host cell messenger RNAs to newly synthesized viral transcripts (Krug *et al.*, 1981), and the similarities in influenza C replication suggest that similar events may be required for the replication of influenza C viruses.

Studies of the morphogenesis of influenza C virus in MDCK cells (Herrler *et al.*, 1981) revealed typical virions with electron-dense nucleocapsids budding at the cell surface. In addition, tightly packed clusters of surface projections were also observed on crescent shaped outfoldings of the plasma membrane on cell surfaces where virus maturation was occurring; such areas were apparently devoid of under-

lying dense nucleocapsid structures. Released particles lacking nucleocapsids were also frequently observed. These findings suggest that in the case of influenza C virus, nucleocapsids may not be required to initiate the budding process at the cell surface, and that interactions between envelope components are sufficient to lead to outfolding of the membrane and particle formation. This is in contrast to observations with influenza A virus, where formation of "empty" particles is rarely detected.

MDCK cells form monolayers in which tight junctions separate the plasma membrane into two domains, and influenza A and parainfluenza viruses are assembled at the free apical surface (Rodriguez Boulan and Sabatini, 1978; Roth *et al.*, 1979). Influenza C virus was similarly found to be assembled at the free apical surface (Herrler *et al.*, 1981). This results indicates that apical virus maturation cannot be directly correlated with the presence or absence of viral neuraminidase or sialic acid in the viral envelope, since influenza C virus contains sialic acid, whereas influenza A and parainfluenza viruses lack this carbohydrate component.

C. Genetics and Epidemiology

Although influenza C virus has only been recognized as the cause of infrequent outbreaks of respiratory illness, a high percentage of individuals possesses antiviral antibody (O'Callaghan *et al.*, 1980). Positive sera were found in 96% of a young adult group, and the highest mean titer was also observed in this age group. Positive sera were less common in young children, and relatively low titers were detected in adults over 65. The high percentage of individuals with antibody to influenza C indicates that infection with this virus occurs at a rate considerably greater than is generally recognized.

Genetic analysis of influenza C virus has been extremely limited, in part because of difficulty in identifying suitable biological or biochemical markers. With the use of plaque morphology and RNA oligonucleotide fingerprints as markers, Racaniello and Palese (1979) have demonstrated that recombination can occur by reassortment of genome segments of influenza C viruses. MDCK cells doubly infected with the C/JHB/66 and C/JJ/50 strains were shown to yield viruses which produced plaques of similar morphology to those of the C/JJ/50 parental type, and RNA fingerprint analysis of one of these isolated clones demonstrated that it contained oligonucleotides from both parents. By analysis of unique oligonucleotide spots that were assigned to individual RNA segments, it was demonstrated that only certain RNA segments from a given parental type were present in the resulting recombinant, indicating that reassortment of genome segments had occurred.

No evidence has been obtained for any antigenic changes in influenza C virus since the first isolation of these agents in 1947. This antigenic stability resembles that of other respiratory agents such as the parainfluenza viruses and respiratory syncytial virus, rather than influenza A or B virus. Nucleotide sequence variation has been examined in 5 representative strains of influenza C virus isolated over a period of 32 years, which were compared by oligonucleotide fingerprint analysis

(Meier-Ewert *et al.*, 1981). Although the oligonucleotide patterns of each strain were found to be distinguishable, the extent of variation was observed to be much less than had been observed in a set of influenza A (H_3N_2) isolates which were isolated over a period of only 5 years (Webster and Laver, 1975). This finding is consistent with the observation that influenza C viruses have not undergone significant antigenic variation. The reason for this difference in stability of the surface antigen is not understood, particularly since influenza C virus appears to share with the influenza A and B viruses most features of virus structure as well as virus replication.

Although it was generally considered that influenza C viruses, like influenza B, occurred only in man, a recent report has described the isolation of influenza C virus from abattoir pigs in Beijing, China (Kuo *et al.*, 1982). Cross neutralization tests and hemagglutination inhibition tests demonstrated that the viruses isolated from pigs were influenza C viruses which were distinguishable from the C/1233/47 reference strain. Nineteen of 100 pigs tested were observed to have significant HI antibody titers to influenza C virus. This report raises the possibility that pigs may serve as natural reservoirs for human influenza C virus.

Acknowledgements

Research by the authors was supported by grants AI 12680, AI 15343, AI 18203, CA 13148 and CA 18611 from the USPHS, PCM 80-6498 from the National Science Foundation, and contract NO1 AI 12677 from the National Institute of Allergy and Infectious Diseases.

References

Air, G. M.: Nucleotide sequence coding for the "signal peptide" and N-terminus of the hemagglutinin from an Asian (H2N2) strain of influenza virus. Virology *97*, 468–472 (1979).

Air, G. M., Hackett, J. A.: Gene 8 of influenza virus: sequences of cDNA transcribed from the 3′ ends of viral RNA of influenza A and B strains. Virology *103*, 291–298 (1980).

Apostolov, K., Flewett, T. H., Kendal, A. P.: Morphology of influenza A, B, C, and infectious bronchitis virus (IBV) virions and their replication. In: The Biology of Large RNA Viruses (Barry, R. D., Mahy, B. W. J., eds.), 3–26. 1970.

Baine, W. B., Luby, J. P., Markin, S. M.: Severe illness with influenza B. Am. J. med. *68*, 181–188 (1980).

Barclay, G. R., Leader-Williams, L. K., Flewett, T. H.: Aspects of influenza C virus replication. J. Hyg. *69*, 587–592 (1971).

Bentley, D. R., Brownlee, G. G.: Sequence of the N2 neuraminidase from influenza virus A/NT/60/68. Nucl. Acids Res. *10*, 5033–5042 (1982).

Blok, J., Air, G. M.: Variation in membrane-insertion and "stalk" sequences in eight subtypes of influenza type A virus neuraminidase. Biochemistry *21*, 4001–4007 (1982 a).

Blok, J., Air, G. M.: Sequence variation at the 3′ end of the neuraminidase gene from 39 type A influenza viruses. Virology *121*, 211–229 (1982 b).

Blok, J., Air, G. M., Laver, W. G., Ward, C. W., Lilley, G. G., Woods, E. F., Roxburgh, C. M., Inglis, A. S.: Studies on the size, chemical composition and partial sequence of the neuraminidase (NA) from type A influenza viruses show that the N-terminal region of the NA is not processed and serves to anchor the NA in the viral membrane. Virology *119*, 109–121 (1982).

Briedis, D. J., Lamb, R. A., Choppin, P. W.,: Influenza B virus RNA segment 8 codes for two nonstructural proteins. Virology *112*, 417–425 (1981).

Briedis, D. J., Lamb, R. A.: The influenza B virus genome: Sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS_1 and NS_2 proteins. J. Virol. *42*, 186–193 (1982).

Briedis, D. J., Lamb, R. A., Choppin, P. W.: Sequence of RNA segment 7 of the influenza B virus genome: partial amino acid homology between the membrane proteins (M1) of influenza A and B viruses and conservation of a second open reading frame. Virology *116,* 581–588 (1982).

Chakraverty, P.: Antigenic relationships between influenza B viruses. Bull. Wld. Hlth. Org. *45,* 755–766 (1972).

Choppin, P. W., Lazarowitz, S. G., Goldberg, A. R.: Studies on proteolytic cleavage and glycosylation of the hemagglutinin of influenza A and B viruses. In: Negative Strand Viruses and Host Cell (Barry, R. D., Mahy, B., eds.), 105–119. 1975.

Clerx, J., Clerx van-Haaster, C., Meier-Ewert, H.: Submitted for publication, 1983.

Compans, R. W., Klenk, H. D., Caliguiri, L. A., Choppin, P. W.: Influenza virus proteins. I. Analysis of polypeptides of the virion and identification of spike glycoproteins. Virology *42,* 880–889 (1970).

Compans, R. W., Content, J., Duesberg, P. H.: Structure of the ribonucleoprotein of influenza virus. J. Virol. *10,* 795–800 (1972).

Compans, R. W., Bishop, D. H. L., Meier-Ewert, H.: Structural components of influenza C virions. J. Virol. *21,* 658–665 (1977).

Corey, L., Rubin, R. J., Thompson, T. R., Noble, G. R., Cassidy, E., Hattwick, M. A. W., Gregg, M. B., Eddins, D.: Influenza B-associated Reye's syndrome: incidence in Michigan and potential of prevention. J. Infect. Dis. *135,* 398–407 (1977).

Desselberger, U., Palese, P.: Molecular weights of RNA segments of influenza A and B viruses. Virology *88,* 394–399 (1978).

Desselberger, U., Racaniello, V. R., Zazra, J. J., Palese, P.: The 3′ and 5′ terminal sequences of influenza A, B, and C virus RNA segments are highly conserved and show partial inverted complementarity. Gene *8,* 315–328 (1980).

Elder, K. T., Bye, J. M., Skehel, J. J., Waterfield, M. D., Smith, A. E.: *In vitro* synthesis, glycosylation and membrane insertion of influenza virus hemagglutinin. Virology *95,* 343–350 (1979).

Fields, S., Winter, G., Brownlee, G. G.: Structure of the neuraminidase gene from human influenza virus A/PR/8/34. Nature (Lond.) *290,* 213–217 (1981).

Flewett, T. H., Apostolov, K.: A reticular structure in the wall of influenza C virus. J. Gen. Virol. *1,* 297–304 (1967).

Gething, M. J., White, J. M., Waterfield, M. D.: Purification of the fusion protein of Sendai virus. Analysis of the NH_2-terminal sequence generated during precursor activation. Proc. Natl. Acad. Sci. U.S.A. *75,* 2737–2740 (1978).

Helenius, A., Kartenbeck, J., Simons, K., Fries, E.: On the entry of Semliki Forest virus into BHK-21 cells. J. Cell Biol. *84,* 404–420 (1980).

Herrler, G., Compans, R. W., Meier-Ewert, H.: A precursor glycoprotein in influenza C virus. Virology *99,* 49–56 (1979).

Herrler, G., Nagele, A., Meier-Ewert, H., Bhown, A. S., Compans, R. W.: Isolation and structural analysis of influenza C virion glycoproteins. Virology *113,* 439–451 (1981).

Hirst, G. K.: The relationship of the receptors of a new strain of virus to those of the mumps-NDV-influenza group. J. Exp. Med. *91,* 177–185 (1950).

Huang, R. T. C., Rott, R., Klenk, H.-D.: Influenza viruses cause hemolysis and fusion of cells. Virology *110,* 243–247 (1981).

Hiti, A. L., Nayak, D. P.: Complete nucleotide sequence of the neuraminidase gene of human influenza virus A/WSN/33. J. Virol. *41,* 730–734 (1982).

Inglis, S. C., Barrett, T., Brown, C. M., Almond, J. W.: The smallest genome RNA segment of influenza virus contains two genes that may overlap. Proc. Nat. Acad. Sci. U.S.A. *76,* 3790–3794 (1979).

Kendal, A. P.: A comparison of influenza C with prototype myxoviruses: Receptor-destroying activity (neuraminidase) and structural polypeptides. Virology *65,* 87–99 (1975).

Klenk, H.-D., Compans, R. W., Choppin, P. W.: An electron microscopic study of the presence or absence of neuraminic acid in enveloped viruses. Virology *42,* 1158–1162 (1970).

Klenk, H.-D., Rott, R., Orlich, M., Blodorn, J.: Activation of influenza A viruses by trypsin treatment. Virology *68,* 426–439 (1975).

Kozak, M.: Mechanism of mRNA recognition by eukaryotic ribosomes during initiation of protein synthesis. Current Topics in Microbiol. and Immunol. *93,* 81–123 (1981).

Krug, R. M., Broni, B. A., Bouloy, M.: Are the 5′ ends of influenza viral mRNAs synthesized *in vivo* donated by host mRNAs? Cell *18*, 329–334 (1979).

Krug, R. M., Plotch, S. J., Ulmanen, I., Bouloy, M., Herz, C., Beaton, A.: The initiation of influenza viral RNA transcription by capped RNA primers. In: Genetic Variation Among Influenza Viruses (Nayak, D. P., ed.), 93–112. New York: Academic Press 1981.

Krystal, M., Elliott, R. M., Benz, E. W., Young, J. F., Palese, P.: Evolution of influenza A and B viruses: conservation of structural features in the hemagglutinin genes. Proc. Natl. Sci. *79*, 4800–4804 (1982).

Kuo, Y. C., Jin, F., Wang, M., Wang, P., Chu, C. M.: Influenza C virus isolated from pigs in China. Kexue Tongbao *27*, 1118–1121 (1982).

Lamb, R. A., Choppin, P. W.: Segment 8 of the influenza virus genome is unique in coding for two polypeptide. Proc. Nat. Acad. Sci. U.S.A. *76*, 4908–4912 (1979).

Lamb, R. A., Lai, C.-J.: Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. Cell *21*, 475–485 (1980).

Lamb, R. A., Choppin, P. W.: Identification of a second protein (M_2) encoded by RNA segment 7 of influenza virus. Virology *112*, 737–745 (1981).

Lamb, R. A., Lai, C.-J.: Conservation of the influenza virus membrane protein (M_1) amino sequence and an open reading frame of RNA segment 7 encoding a second protein (M_2) in H_1N_1 and H_3N_2 strains. Virology *112*, 746–751 (1981).

Lamb, R. A., Lai, C.-J., Choppin, P. W.: Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: Colinear and interrupted mRNAs code for overlapping proteins. Proc. Nat. Acad. Sci. U.S.A. *78*, 4170–4174 (1981).

Laver, W. G., Baker, N.: Amino acid composition of polypeptides from influenza virus particles. J. Gen. Virol. *17*, 61–67 (1972).

Lazarowitz, S. G., Choppin, P. W.: Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. Virology *68*, 440–454 (1975).

Lenard, J., Miller, D. K.: PH-dependent hemolysis by influenza, Semliki Forest virus, and Sendai virus. Virology *110*, 479–482 (1981).

Lu, B. L., Webster, R. G., Brown, L. E., Nerome, K.: Analysis of antigenic drift in influenza B viruses: do pooled monoclonal antibodies have any utility as a standard reference reagent? Bull. WHO (in press).

Maeda, T., Ohnishi, S.: Activation of influenza virus by acidic media causes hemolysis and fusion of erythrocytes. FEBS Lett. *122*, 283–287 (1980).

Markoff, L., Lai, C.-J.: Sequence of the influenza A/Udorn/72 (H3N2) virus neuraminidase gene as determined from cloned full-length DNA. Virology *119*, 288–297 (1982).

Martin, M. L., Palmer, E. L., Kendal, A. P.: Lack of characteristic hexagonal surface structure on a newly isolated influenza C virus. J. Clinical Micro. *6*, 84–86 (1977).

Meier-Ewert, H., Compans, R. W., Bishop, D. H. L., Herrler, G.: Molecular analysis of influenza C virus. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 127–133. New York: Academic Press 1978.

Meier-Ewert, H., Herrler, G., Nagele, A., Compans, R. W.: Glycoproteins of influenza C virus. In: Structure and Variation in Influenza Virus (Laver, G., Air, G., eds.), Vol. 5, 357–366. Elsevier/North-Holland 1980.

Meier-Ewert, H., Nagele, A., Herrler, G., Basak, S., Compans, R. W.: Analysis of influenza C virus structural proteins and identification of a virion RNA polymerase. In: Replication of Negative Strand Viruses (Bishop, D. H. L., Compans, R. W., eds.), 173–180. New York: Elsevier 1981 a.

Meier-Ewert, H., Petri, T., Bishop, D. H. L.: Oligonucleotide fingerprint analyses of influenza C virion RNA recovered from five different isolates. Arch. Virol. *67*, 141–147 (1981 b).

Min Jou, W., Verhoeyen, M., Devos, R., Saman, E., Fang, R., Huylebroeck, D., Fiers, W., Threlfall, G., Barber, C., Carey, N., Emtage, S.: Complete structure of the hemagglutinin gene from the human influenza A/Victoria/3/75 (H3N2) strain as determined from cloned DNA. Cell *19*, 686–696 (1980).

Morbidity and Mortality Weekly Report: Follow-up on Reye Syndrome *29*, 321–322 (1980 a).

Morbidity and Mortality Weekly Report: Reye Syndrome Ohio, Michigan *29*, 532–538 (1980 b).

Morbidity and Mortality Weekly Report: Influenza Worldwide *31*, 494–495 (1982).

Nagele, A., Meier-Ewert, H.: Submitted for publication, 1983.
Nakamura, K., Compans, R. W.: Glycopeptide components of influenza viral glycoproteins. Virology *86,* 432–442 (1978).
Nakamura, K., Compans, R. W.: Biosynthesis of the oligosaccharides of influenza virus glycoproteins. Virology *93,* 31–47 (1979).
Nakamura, K., Herrler, G., Petri, T., Meier-Ewert, H., Compans, R. W.: Carbohydrate components of influenza C virions. J. Virol. *29,* 997–1005 (1979).
Nelson, D. B., Hurwitz, E. S., Sullivan-Bolyai, J. A., Morens, D. M., Schonberger, L. B.: Reye's syndrome in the United States in 1977–1978, a non-influenza year. J. Infect. Dis. *140,* 436–439 (1979).
Nerome, K., Ishida, M.: The multiplication of an influenza C virus in an established line of canine kidney (MDCK) cells. J. Gen. Virol. *39,* 179–181 (1978).
Nerome, K., Ishida, M., Nakayama, M.: Established cell line sensitive to influenza C virus. J. Gen. Virol. *43,* 257–259 (1979).
Norman, M. G., Lowden, J. A., Hill, D. E., Bannatyre, R. M.: Encephalopathy and fatty degeneration of the viscera in childhood. Can. Med. Assoc. J. *99,* 549–554 (1968).
O'Callaghan, R. J., Loughlin, S. M., Labat, D. D., Howe, C.: Properties of influenza C virus grown in cell culture. J. Virol. *24,* 875–882 (1977).
O'Callaghan, R. J., Gohd, R. S., Labat, D. D.: Human antibody to influenza C virus: Its age-related distribution and distinction from receptor analogs. Infect. and Immun. *30,* 500–505 (1980).
Ohuchi, M., Homma, M., Muramatsu, M., Ohyama, S.: Properties of the erythrocyte receptors for influenza C virus. Microbiol. Immunol. *22,* 197–203 (1978).
Ohuchi, M., Ohuchi, R., Mifune, K.: Demonstration of hemolytic and fusion activities of influenza C virus. J. Virol. *42,* 1076–1079 (1982).
Oxford, J. S.: Polypeptide composition of influenza B viruses and enzymes associated with the purified virus particles. J. Virol. *12,* 827–835 (1973).
Pereira, H. G.: Influenza: antigenic spectrum. Prog. Med. Virol. *11,* 46–79 (1969).
Pereira, M. S.: The effect of shifts and drifts on the epidemiology of influenza in man. Phil. Trans. Roy. Soc. (Lond.) *B 288,* 423–432 (1980).
Petri, T., Meier-Ewert, H., Compans, R. W.: Replication and plaque assay of influenza C virus in chicken kidney cells. FEMS Microbiol. Letters *5,* 227–230 (1979 a).
Petri, T., Meier-Ewert, H., Compans, R. W.: Inhibition of influenza C virus replication by Actinomycin D, α-amanitin, and UV irradiation. J. Virol. *32,* 1037–1040 (1979 b).
Petri, T., Meier-Ewert, H., Crumpton, W. M., Dimmock, N. J.: RNAs of influenza C virus strains. Arch. Virol. *61,* 239–243 (1979 c).
Petri, T., Herrler, G., Compans, R. W., Meier-Ewert, H.: Gene products of influenza C virus. FEMS Microbiol. Letters *9,* 43–47 (1980).
Pons, M. W., Schulze, I. T., Hirst, G. K.: Isolation and characterization of the ribonucleoprotein of influenza virus. Virology *39,* 250–259 (1969).
Portner, A., Webster, R. G., Bean, W. J.: Similar frequencies of antigenic variants in Sendai, vesicular stomatitis, and influenza A viruses. Virology *104,* 235–238 (1980).
Racaniello, V. R., Palese, P.: Isolation of influenza C virus recombinants. J. Virol. *32,* 1006–1014 (1979).
Racaniello, V. R., Palese, P.: Influenza B virus genome: assignment of viral polypeptides to RNA segments. J. Virol. *29,* 361–373 (1979).
Reye, R. D. K., Morgan, G., Barol, J.: Encephalopathy and fatty degeneration of the viscera: a disease entity in childhood. Lancet *2,* 749–752 (1963).
Ritchey, M. B., Palese, P., Kilbourne, E. D.: RNAs of influenza A, B, and C viruses. J. Virol. *18,* 738–744 (1976).
Robertson, J. S., Schubert, M., Lazzarini, R. A.: Polyadenylation sites for influenza virus mRNA. J. Virol. *38,* 157–163 (1981).
Rodriguez Boulan, E., Sabatini, D. D.: Asymmetric budding of viruses in epithelial monolayers: a model system for study of epithelial polarity. Proc. Natl. Acad. Sci. U.S.A. *74,* 5071–5075 (1978).
Roth, M. G., Fitzpatrick, J. P., Compans, R. W.: Polarity of influenza and vesicular stomatitis virus maturation in MDCK cells: lack of a requirement for glycosylation of viral glycoproteins. Proc. Natl. Acad. Sci. U.S.A. *76,* 6430–6434 (1979).

Scheid, A., Graves, M. C., Silver, S. M., Choppin, P. W.: Studies on the structure and functions of paramyxovirus glycoproteins. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 181–193. London: Academic Press 1978.

Schiff, G. M.: Reye's syndrome. Ann. Rev. Medicine *27,* 447–452 (1976).

Schild, G. C., Pereira, M. S., Chakraverty, P., Coleman, M. T., Dowdle, W. R., Change, W. K.: Antigenic variants of influenza B virus. Brit. Med. J. *4,* 127–131 (1973).

Scholtissek, C., Rohde, W., Harms, E.: Genetic relationship between an influenza A and a B virus. J. Gen. Virol. *37,* 243–247 (1977).

Schwarz, R. T., Schmidt, M. F. G., Anwer, U., Klenk, H.-D.: Carbohydrates of influenza virus. I. Glycopeptides derived from viral glycoproteins after labeling with radioactive sugars. J. Virol. *23,* 217–226 (1977).

Shaw, M. W., Lamb, R. A., Erickson, B. W., Briedis, D. J., Choppin, P. W.: Complete nucleotide sequence of the neuraminidase gene of influenza B virus. Proc. Natl. Acad. Sci. *79,* 6817–6821 (1982).

Skehel, J. J., Hay, A. J.: Nucleotide sequences at the 5′ termini of influenza virus RNAs and their transcripts. Nucleic Acids Res. *5,* 1207–1219 (1978).

Stuart-Harris, C. H., Schild, G. C.: In: Influenza: The Viruses and the Disease, 22–27. London: Edward Arnold 1976.

Sugawara, K., Ohuchi, M., Nakamura, K., Homma, M.: Effects of various proteases on the glycoprotein composition and the infectivity of influenza C virus. Arch. Virol. *68,* 147–151 (1981).

Tobita, K., Kilbourne, E. D.: Genetic recombination for antigenic markers of antigenically different strains of influenza B virus. J. Virol. *13,* 347–352 (1974).

Ueda, M., Tobita, A., Suguira, A., Enomoto, C.: Identification of hemagglutinin and neuraminidase genes of influenza B virus. J. Virol. *25,* 685–686 (1978).

Verhoeyen, M., Fang, R., Min Jou, W., Devos, R., Huylebroeck, D., Saman, E., Fiers, W.: Antigenic drift between the hemagglutinin of the Hong Kong influenza strains A/Aichi/2/68 and A/Victoria/3/75. Nature (Lond.) *286,* 771–776 (1980).

Ward, C. W.: Structure of the influenza virus hemagglutinin. Curr. Topics Microbiol. and Immunol. *94/95,* 1–74 (1981).

Waterfield, M. D., Espelie, K., Skehel, J. J.: Structure of the hemagglutinin of influenza virus. Br. Med. Bull. *35,* 57–63 (1979).

Waterson, A. P., Hurrell, J. M. W., Jensen, K. E.: The fine structure of influenza A, B, and C viruses. Arch. ges. Virusforsch. *12,* 487–495 (1963).

Webster, R. G., Berton, M. T.: Analysis of antigenic drift in the hemagglutinin molecule of influenza B viruses with monoclonal antibodies. J. Gen. Virol. *54,* 243–251 (1981).

Webster, R. G., Hinshaw, V. S.: Matrix protein from influenza A virus and its role in cross-protection in mice. Infect. Immun. *17,* 561–566 (1977).

Webster, R. G., Laver, W. G.: Antigenic variation of influenza viruses. In: The Influenza Viruses and Influenza (Kilbourne, E. D., ed.), 269–314. New York: Academic Press 1975.

Wilson, I. A., Skehel, J. J., Wiley, D. C.: Structure of the hemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature (Lond.) *289,* 366–373 (1981).

Winter, G., Fields, S.: Cloning of influenza DNA into M13: The sequence of the RNA segment encoding the A/PR/8/34 matrix protein. Nucleic Acids Res. *8,* 1965–1974 (1980).

Winter, G., Fields, S., Brownlee, G. G.: Nucleotide sequence of the hemagglutinin gene of a human influenza virus H1 subtype. Nature (Lond.) *292,* 72–75 (1981).

10

Virus-Determined Differences in the Pathogenesis of Influenza Virus Infections

J. L. Schulman

I. Introduction

Among the variations in phenotype expressed by different influenza viruses are differences in biological activity in various host cell systems, including differences in host range, organ specificity and dissemination in animals and plaquing efficiency in various cell culture systems. With the availability of reliable laboratory techniques to determine the parental origins of individual RNA segments in reassortant viruses obtained in nature or in the laboratory, it has become possible to obtain previously unavailable information concerning the contribution of individual genes or combinations of genes to some of these biological differences and to begin to address some fundamental questions regarding the molecular mechanisms involved.

The discussion which follows attempts to review currently available information obtained from comparisons of wild type and reassortant viruses in different host cell systems. It should be stated at the outset that this review will concentrate on the genetic basis for biological differences among different strains of influenza A virus. No attempt will be made to address very important questions regarding the basis of the relatively restricted host ranges of influenza B and C viruses in nature, simply because of the paucity of relevant data.

It should also be understood that while the review concentrates on virus-determined differences in biological activity, "virulence" obviously reflects a complex set of interactions in which host-determined differences are also of crucial importance. For example, influenza viruses prevalent in avian populations, rarely (if ever) directly cross species barriers to infect humans (reviewed by Hinshaw and Webster, 1982) and avian influenza viruses which are lethal in turkeys and chickens do not produce disease in other avian species (Slemons and Easterday, 1972; Alexander *et al.*, 1978). With a few notable exceptions (discussed below), only limited information concerning the mechanism of host-determined differences in susceptibility is available. For example, Haller *et al.* (1979) have demonstrated that the genetically determined resistance of A2G mice bearing the Mx gene is mediated

by differential responses to interferon. In other studies, it has been clearly established that lysis of influenza virus-infected target cells *in vitro* by human or murine cytolytic T cells requires recognition of virus-encoded antigens in association with MHC-encoded cell surface antigens, and evidence has been obtained indicating: a) that the T cell response may be influenced by other MHC gene products and b) that different HLA antigens may be used preferentially during lysis of influenza A virus- and influenza B virus-infected targets (reviewed by Askonas *et al.*, 1982). Nevertheless, it should be emphasized that the precise relationship between susceptibility to influenza virus infection under natural conditions and genetically controlled immune responses to influenza virus antigens remains undefined.

Finally, it should be stressed that no effort will be made to comprehensively review the critical role of specific immunity in influencing viral pathogenesis *in vivo*. It has been shown repeatedly that specific antibody to hemagglutinin can prevent infection or significantly modify infection and that antibody to neuraminidase also provides protection (reviewed by Schulman, 1975). More recently, similar protective effects have been demonstrated following the parenteral administration to mice of monoclonal antibodies specific for hemagglutinin or neuraminidase (Askonas *et al.*, 1982). In addition, results of recent studies of cellular immune responses in mice and in humans suggest that cellular immunity may play a critical role during the recovery period, when virus is cleared from infected tissues. Furthermore, these cellular immune responses are characterized by much more cross-reactivity among influenza A virus subtypes than can be detected by conventional serologic assay (reviewed by Askonas *et al.*, 1982). Consequently, it is clear that meaningful *in vivo* comparisons of the virulence of different influenza A virus isolates can only be made in hosts not previously primed by prior immunologic experience with related influenza viruses.

II. Methods

Correlations of specific biological phenotypes with specific genes or gene products of influenza viruses have relied on comparisons of closely related wild type or mutant strains or on the analysis of reassortant viruses in which the parental derivation of each of the 8 RNA segments has been defined. In the former instance, comparisons of different wild type strains with respect to replication in different host cell systems have been related to differences in a particular gene product. Most notably, studies of different avian influenza viruses have clearly demonstrated that pathogenicity for chickens is related to susceptibility of the hemagglutinin to proteolytic cleavage in different host cell systems (Bosch *et al.*, 1979). In particular, it

was shown in chorioallantoic membrane cultures that avirulent strains could replicate productively in allantoic epithelial cells but that progeny produced in chorionic cells lacked a cleaved hemagglutinin and were non-infectious. In contrast, the hemagglutinin of virulent strains was cleaved in both cell types (Rott *et al.*, 1980).

In the second system, reassortant viruses obtained after mixed infection with parental viruses with readily distinguishable phenotypes are analyzed to identify the parental origin of each RNA segment and then are assayed in the appropriate host cell systems to determine which RNA segment(s) of the virulent parent is/are required for virulence. The methods employed in genotyping reassortant viruses, described in greater detail in Chapter 4, include comparisons of parental and reassortant viruses with respect to: migrational patterns of RNAs on polyacrylamide gels (Palese and Schulman, 1976); protein patterns on polyacrylamide gels (Ritchey *et al.*, 1977); oligonucleotide fingerprinting of individual RNA segments (Young and Palese, 1979); direct molecular hybridization of cRNA to radiolabelled individual virion RNA segments (Scholtissek *et al.*, 1976) and by competitive hybridization in which RNA segments of test viruses are used to compete for binding of labelled vRNA to cRNA (Bean *et al.*, 1980).

With rare exceptions, these techniques provide reliable methods by which laboratory-derived reassortant viruses can be genotyped. However, when applied to the analysis of viruses isolated in nature, the data which are obtained are generally not as conclusive. Since precise parental strains are unknown, it is possible only to determine which prototype virus contains an RNA segment most closely related to a particular segment in a putative reassortant virus.

More importantly, during the isolation and purification of reassortant viruses there are random mutational events which may influence the phenotype (Scholtissek *et al.*, 1977; Schulman and Palese, 1978; Erickson and Kilbourne, 1980; Ogawa and Ueda, 1981; Campbell *et al.*, 1982). Similar observations of reassortant viruses have been made in the reovirus system, where it has been shown that neurovirulence requires the S1 gene of type 3 reovirus. However, following the identification of a non-neurovirulent type 3 mutant, it was demonstrated that neurovirulence was restored in reassortant viruses deriving the M2 gene from a different parent (Hrdy *et al.*, 1982). Thus, meaningful correlation of gene constellation with biologic properties depends on examination of several reassortant viruses of identical genotype.

In addition, the phenotypes of reassortant viruses may depend in part on effective interaction of particular gene products. Lubeck *et al.* (1979) demonstrated that genetic reassortment of influenza A viruses is not completely random in that certain combinations of RNA segments derived from a given parent were observed much more frequently than expected. The association was most striking with respect to genes coding for the P proteins. In addition, Rott *et al.* (1979) found that among reassortant viruses obtained after mixed infection with two virulent avian viruses, with only one exception, reassortant viruses which derived all of the genes coding for the P and NP proteins from the same parent were virulent for chickens. All reassortant viruses in which the derivation of these genes was mixed lacked virulence.

III. Hemagglutinin-Related Differences in Biologic Properties

There is no doubt that at present the most clearly defined virus-determined difference in virulence is related to cleavage of the hemagglutinin. Klenk *et al.* (1975) and Lazarowitz and Choppin (1975) first demonstrated that virus particles produced in cells which lack host proteases capable of cleaving viral hemagglutinin into HA1 and HA2 subunits can attach to receptors on host cells, but are incapable of initiating infection. Since then, it has been shown in a variety of cell systems and in intact animals that the capacities of different influenza viruses to undergo multicycle replication requires that the host cell employed can carry out the necessary cleavage event (Bosch *et al.*, 1979). It has since been shown that this property is not strictly associated with the hemagglutinin serotype, in that some H7-containing viruses are pathogenic and replicate with a cleaved HA, whereas the hemagglutinins of other avirulent viruses of the same serotype are not cleaved (reviewed by Hinshaw and Webster, 1982). Furthermore, these differences have been correlated with sequence data on carboxyl termini of different HA1 species, which demonstrate that in H2- and H3-containing human viruses, and in avirulent H7-containing avian viruses, a single arginine-glycine peptide is susceptible to proteolytic cleavage, whereas the connecting peptide at the carboxyl terminus of HA1 of virulent H7 strains is longer and contains basic peptides. Accordingly, it has been proposed that these structural differences at the cleavage site are directly related to the "cleavability" of different hemagglutinins in different host cell systems (Bosch *et al.*, 1981).

While it is clear from the above results that susceptibility of the hemagglutinins of avian influenza viruses to cleavage is determined by the structure of the hemagglutinin, in other systems differences in viral neuraminidase may determine whether or not cleavage of HA occurs. Influenza A/WSN/33 (H1N1) virus is unique among human influenza viruses in its capacity to undergo multicycle replication and to produce plaques in MDBK cells. Analysis of reassortant viruses obtained after mixed infection with WSN virus and influenza A/HK/68 (H3N2) or influenza A/FM/1/47 (H1N1) viruses demonstrated that the WSN neuraminidase was required for plaque formation and that reassortant viruses deriving only the neuraminidase gene from WSN virus were capable of producing plaques. Furthermore, it was shown that reassortant viruses lacking the WSN neuraminidase produced non-infectious progeny with uncleaved HA in MDBK cells and that infectivity could be activated by *in vitro* treatment with trypsin (Schulman and Palese, 1976). In related studies it was shown that among reassortant viruses of A/Turkey/Ontario/7732/66 (H5N9) and A/WSN/33 (H1N1) viruses only those which derived the HA gene from WSN virus and the neuraminidase gene from the avian virus failed to produce plaques on CEF cells in the absence of external protease. Furthermore, it was shown that these same reassortant viruses were completely lacking in virulence for 1-day-old chicks (Bean and Webster, 1978). It has also been shown that the neurovirulence of WSN virus in mice was absent in reassortant viruses lacking the WSN neuraminidase (Sugiura and Ueda, 1980) and

that only reassortant viruses containing the WSN neuraminidase contained a cleaved HA and could replicate productively in a neuroblastoma cell line (Nakajima and Suguira, 1980). The unique capacity among human influenza viruses of WSN virus and reassortant viruses containing WSN neuraminidase to replicate with a cleaved HA has also been observed in BHK cells, KB cells, L cells and W1-38 cells (Schulman, unpublished observations). The mechanism by which viral neuraminidase in some instances may facilitate cleavage of HA has not been elucidated, but it is possible that different viral neuraminidases may differentially activate or inactivate the required host proteases.

On the basis of observations of cleavage activation of infectivity of influenza and paramyxoviruses, Scheid *et al.* (1978) proposed a functional analogy between the N-terminal region of the F1 protein of paramyxoviruses and the N-terminus of HA2 of influenza viruses and suggested that both were required for viral penetration. Recent evidence demonstrating fusion of influenza virus and cell membranes at low pH (White *et al.*, 1981; Lenard and Miller, 1981; Huang *et al.*, 1981; White *et al.*, 1982) and the elucidation of common amino acid sequences at the N- termini of these polypeptides (Skehel and Waterfield, 1975; Gething *et al.*, 1978) are in accord with such a hypothesis. Furthermore, it has been shown that liposomes constructed with uncleaved HA are incapable of fusing to cells, whereas, following cleavage of the HA molecules by *in vitro* treatment with trypsin, fusion was observed (Huang *et al.*, 1980). Thus, it is extremely likely that the requirement for proteolytic cleavage of HA for infectivity is directly related to the role of HA2 in promoting fusion of the virus membrane with the host cell membrane, probably within endosomes (White *et al.*, 1982).

However, it should also be pointed out that in addition to this clearly defined relationship between HA cleavage and infectivity, other differences in influenza virus hemagglutinins may also play a role in determining host range. Thus, differences in the avidity with which particular hemagglutinins attach to specific sialic acid receptors on different host cells or differential binding to specific glycoprotein inhibitors on mucus-covered surfaces could be major factors affecting host range.

IV. Virulence in Chickens

The pathogenesis of avian influenza virus infection in avian species is extremely variable, ranging from asymptomatic infection of the gastrointestinal tract of waterbirds to fatal disseminated infection with some viruses in chickens and turkeys (reviewed by Hinshaw and Webster, 1982). The viruses which have been isolated from the gastrointestinal tract or feces of ducks include a great variety of hemagglutinin and neuraminidase subtypes and at present there is insufficient data to link adaptation to gastrointestinal tissues to particular influenza virus genes or constellations of genes.

Much more information has been gathered with respect to RNA segments involved in highly virulent infections. As noted above, a clear correlation between "cleavability" of the hemagglutinin in different cells systems in culture and patho-

genicity for chickens has been established (reviewed by Rott, 1980; 1982). Later, it was shown that parental viruses and reassortant viruses pathogenic for chickens replicated efficiently at 41 °C. In contrast, non-pathogenic viruses containing a readily "cleavable" HA replicated less efficiently at 41 °C, demonstrating that temperature-dependent replication is an additional phenotype of avirulent reassortant viruses (Rott *et al.*, 1982).

Employing *ts* mutants of the highly virulent fowl plague virus (FPV) in crosses with different human and animal influenza viruses, Scholtissek *et al.* (1977) studied virulence for chickens of reassortant viruses in which only one RNA segment was not derived from the FPV parent. Loss of pathogenicity was observed for each FPV segment exchanged, depending on the other parental virus employed. For example, replacement of RNA segment 2 with the corresponding RNA of A/Singapore/57 (H2N2) virus resulted in reassortant viruses as virulent as FPV, whereas replacement of the same segment with the corresponding RNA segment from A/Hong Kong/1/68 (H3N2) virus produced avirulent reassortant viruses. This observation is particularly noteworthy because of the close sequence homology of RNA 2 in Singapore and Hong Kong viruses (Scholtissek *et al.*, 1978) and reinforces the authors' conclusion that the loss of pathogenicity is highly dependent on the particular strain used to replace a specific RNA segment. It should also be noted that reassortants in which the FPV HA gene was exchanged were not studied. Undoubtedly, all of these would have been avirulent because of the absence of the "cleavable" FPV HA. In addition, reassortants deriving the FPV HA gene and RNA 7 (coding for M1 and M2 polypeptides) from the other parent were not available, so it is not possible to determine whether or not all reassortants lacking RNA 7 of FPV would be avirulent, regardless of which strain was employed as a donor.

In another study, reassortants obtained after mixed infection with A/Turkey/Ontario 7732/66 (H5N9) and A/WSN/33 (H1N1) were assayed for pathogenicity in 1-day-old chicks. Only those reassortant viruses deriving the turkey virus HA or WSN virus NA (or both) produced plaques on chick embryo fibroblast cells and were virulent in chicks. In addition, it was found that all virulent reassortant viruses derived both RNA 3 (most likely coding for PA) and RNA 5 (coding for NP) from the virulent turkey virus parent. Since reassortants in which these 2 segments were segregated were not obtained, it is not possible to determine whether virulence requires that only one or both of these segments be derived from the virulent parent. Ogawa and Ueda (1981) examined reassortants derived from A/chicken/Japan/24 (H7N7) and A/duck/Ukraine/1/63 (H3N8) viruses, and found that reassortants deriving the HA gene, the NA gene and segment 7 (coding for M1 and M2 polypeptides) from the virulent A/Japan virus parent were as virulent as the virulent parental strain in chick embryos and 10-day-old chickens. Reassortants lacking the H7 HA were completely avirulent, whereas reassortants containing the H7 HA which did not derive either the NA gene or RNA 7 from the virulent parent had diminished virulence. Nevertheless, they were clearly more pathogenic than the avirulent A/duck/Ukraine parent. The N7 NA of the virulent Japan virus had no effect on cleavage of the H7 HA in cell culture, but its contribution to virulence was confirmed by obtaining H7N7 reassortants following mixed infection with H7N8 and H3N7 antigenic hybrids.

Clearly, the one common feature in all these studies is the association of cleavage of HA with virulence. With respect to the contributions of other RNA segments to virulence, it appears equally clear that virulence for chickens is polygenic and that the importance of particular RNA segments depends upon the parental strains employed.

V. Neurovirulence in Mice

The unique capacity of some strains of influenza virus to replicate in mouse neuronal cells has been employed as a genetic marker for 3 decades by scientists studying the genetics of influenza viruses (reviewed by Sugiura, 1975). Retrospectively, evidence linking neurovirulence with neuraminidase activity was obtained in early experiments (Burnet, 1959). Later, in studies of antigenic hybrids of A/NWS/33 (H1N1) and A/Japan/305/57 (H2N2) viruses, it was concluded that neurovirulence is not exclusively linked to either hemagglutinin or neuraminidase (Mayer *et al.*, 1973). However, further examination of genotyped reassortant viruses derived from WSN virus demonstrated that WSN neuraminidase was required for neurovirulence (Sugiura and Ueda, 1980). In addition, these investigators found that only those reassortant viruses which also derived RNA segments 7 and 8 from WSN virus were as neurovirulent as WSN virus. Viruses which contained WSN virus NA but which derived either of those other RNA segments from the avirulent HK virus parent replicated to much lower titer in mouse brains and were not lethal except in immunosuppressed mice. In related studies, in which replication of the same viruses was studied in a neuroblastoma cell line, it was demonstrated that only reassortant viruses containing the WSN virus NA replicated with a cleaved HA in these cells. (Nakajima and Sugiura, 1980).

Scholtissek *et al.* (1979) examined the neurovirulence of reassortant viruses derived from A/FPV/Rostock/34 (H7N1) and A/PR/8/34 (H1N1) or A/England/1/61 (H2N2) viruses. Although none of the parental strains is neurovirulent in mice, neurovirulent reassortant viruses were identified. As summarized in Table 1, neurovirulent reassortants obtained after mixed infection with FPV and H2N2 virus derived RNA segments 1 and 2 (coding for PB2 and PB1, respectively) from the H2N2 virus parent and RNA segments 4 and 7 (coding for HA, and M1 and M2, respectively) from FPV. When PR8 virus was employed as the other parent, neurovirulence was associated with the PR8 virus gene for PB1 and FPV segments 4 and 7. Moreover, in related studies it was shown that the same reassortants which were neurovirulent following intracerebral inoculation of suckling mice were also capable of dissemination with viremia and CNS involvement following infection of the respiratory tract. In addition, it was noted that only those reassortants capable of dissemination *in vivo* were capable of multicycle replication in mouse embryo cells in culture. Viruses containing H1 or H2 hemagglutinin which were pneumotropic but not capable of dissemination *in vivo* could replicate in mouse embryo cells in culture in the presence of trypsin (Vallbrecht *et al.*, 1980). These observations clearly demonstrate that virulent reassortants can be derived

following mixed infection with 2 avirulent parental viruses, and once again confirm the critical role of the "cleavable" H7 HA in dissemination of infection.

VI. Virulence in Other Experimental Animals

A. Ferrets

Ferrets have been employed for several decades as a useful animal model of influenza, particularly with respect to comparisons of virulence among different influenza virus isolates, with the aim of predicting virulence for man. In recent years, some of these experiments have included comparisons of reassortant viruses of defined genotype. For example, Campbell *et al.* (1982) compared 16 reassortant viruses derived from A/Okuda/57 (H2N2) and A/Finland/4/74 (H3N2) viruses and found that those deriving genes coding for HA and NA from the attenuated Okuda virus parent tended to be less virulent than those which derived these RNA segments from the virulent parent. Although virulence could not be associated with a specific gene constellation, an inverse relationship was noted between the number of RNA segments derived from the avirulent parent and the virulence of reassortant viruses.

B. Rats

Similar experiments employing both wild type and reassortant viruses have been conducted in infant rats. For some of the viruses, information concerning virulence in human volunteers was also avialable and in general there appeared to be some correlation between pathogenicity in infant rats as determined by replication in nasal turbinates and susceptibility to superinfection with *Haemophilus Influenzae B* and virulence in human subjects. However, an insufficient number of reassortants were examined to permit discrimination of particular RNA segments or constellations of segments which contribute to pathogenicity (Michaels *et al.*, 1978; Teh *et al.*, 1980; Jennings *et al.*, 1980).

C. Mice

In their studies of neurovirulence in mice, of reassortant viruses obtained from FPV and mouse-adapted H2N2 or H1N1 viruses, Scholtissek *et al.* (1979) also examined the pneumovirulence of the same viruses following intranasal inoculation of suckling mice. As summarized in Table 1, pneumovirulence required only that the RNA segment coding for PA was derived from the mouse pneumotropic influenza virus parent. In addition, it was noted that all pneumotropic viruses could replicate in mouse kidney cells in culture. One reassortant virus deriving the segment coding for PA from A/England/61 (H2N2) virus lacked pneumovirulence and did not replicate productively in mouse kidney cells.

In other unpublished studies, the pneumovirulence of antigenic hybrid viruses obtained from mouse adapted A/PR8 (H1N1) and non-adapted A/HK/8/68 (H3N2) viruses was compared. As shown in Table 2, a reassortant virus (R1), deriving all of

Table 1. *Gene constellations required for neurovirulence or pneumovirulence in mice among reassortant viruses derived from FPV and A/England/1/61 or FPV and A/PR/8/34 viruses*

	FPV × A/Eng.		FPV × PR8	
	Neurovirulence	Pneumnovirulence	Neurovirulence	Pneumnovirulence
PB1	E	–	–	–
PA	E	E	P	P
HA	F	–	F	–
$M_{1,2}$	F	–	F	–
all other genes	–	–	–	–

Data adapted from Scholtissek *et al.*, 1979.
E = RNA segment derived from A/Eng virus.
P = RNA segment derived from A/PR/8 virus.
F = RNA segment derived from FPV.
– = RNA segment derived from either parent.

its RNA segments from PR8 virus except the HA gene, was less virulent than its PR8 virus parent. Conversely, R2, which derived only its HA gene from PR8 virus, was more virulent than its HK virus parent from which all of its other genes were derived. (Similar results were obtained with other reassortant viruses identical in genotype to R1 or R2.) Following mixed infection with these 2 reassortants, R3, deriving all of its genes from PR8 was obtained, and its virulence for mice was equivalent to that of the PR8 virus wild type and greater than that of either R1 or R2. These results indicate once again that the HA is a major factor in determining virulence. However, it is not known whether this HA-dependent difference in virulence is mediated by differences in the efficiency of cleavage in cells of the mouse respiratory tract. These observations also suggest that other genes also contribute to the virulence of PR8 virus. Thus, in related studies, a reassortant mixture obtained following mixed infection with PR8 and HK viruses was inoculated undiluted into mice. Viruses taken from infected mouse lungs were labelled in MDCK cells to determine which RNA segments were present as doublets, indicating derivation from both parents, and which were present as single bands, indicating selection for

Table 2. *Pneumovirulence of reassortant viruses of A/PR/8/34 and A/HK/68 viruses*

	Genotype							Virus titers		Lung
Virus	PB1	PA	PB2	HA	NP	NA	M1, 2	NS1, 2	in lungs[b]	lesions[c]
PR8	P[a]	P	P	P	P	P	P	P	7.7	73
HK	H[a]	H	H	H	H	H	H	H	6.2	5
R1	P	P	P	H	P	P	P	P	6.9	33
R2	H	H	H	P	H	H	H	H	6.8	28
R3[d]	P	P	P	P	P	P	P	P	7.9	68

[a] P = PR8 virus derived; H = HK virus derived.
[b] EID_{50}, log_{10}, geometric mean 3 days after infection.
[c] mean extent of lung lesions (%) 7 days after infection.
[d] obtained following mixed infection with R1 and R2.

reassortants which derived those RNA segments from PR8 virus. Only segment 8 (coding for NS1 and NS2) was present as a doublet. HK virus segments coding for PB1, PA, PB2, NP or M1 and M2 were not detected, indicating selection for PR8 virus-derived genes for these proteins during replication in mouse lungs (Schulman, J. L., previously unpublished). In contrast, in earlier studies when the same mixture was inoculated into eggs, selection for PR8 virus genes coding for P proteins or NS1 or NS2 proteins was not observed (Schulman and Palese, 1978). Thus, it is likely that during adaptation of PR8 virus to mice by multiple passage, selection has taken place for changes in most RNA segments.

Hepatotropism: Haller (1975) derived an avian influenza A virus with the unique capacity to replicate in mouse liver following intraperitoneal inoculation or infection of the respiratory tract. Reassortants derived from that virus and A/PR/8 virus were tested for hepatotropism, and, as summarized in Table 3, hepatotropism was found to require the presence of RNA segments coding for PA, HA, NP and M1 and M2 from the hepatotropic virus parent (Schulman, J. L., previously unpublished data).

VII. Virus-Determined Differences in Other Systems

A. Cell Culture

The limited capacity of FPV to replicate productively in some cells of mammalian origin has been associated both with lack of migration of RNP from the nucleus and with decreased secondary synthesis of membrane protein (reviewed by Conti *et al.*, 1980). In two different cell culture systems, the gene coding for PB2 has been shown to control host range. Thus, Almond (1977) demonstrated that the PB2 gene of FPV Dobson determined its ability to undergo multicycle replication in BHK cells and Israel (1980) found that the same gene determined the capacity of a mutant strain of FPV Dobson to replicate in L cells.

Table 3. *Hepatotropic and non-hepatotropic reassortant viruses*

PB1	PA	PB2	HA	NP	NA	M1, 2	NS1, 2	Hepatotropism
P	T	T	T	T	P	T	P	+
T	T	P	T	T	T	T	T	+
T	T	P	T	T	P	T	P	+
P	T	T	T	T	P	T	T	+
P	P	P	T	T	T	T	T	−
T	T	T	P	T	T	T	T	−
P	P	P	T	P	T	P	T	−
T	T	T	T	T	T	P	T	−

P = PR8 virus derived.
T = Turkey virus derived.

B. Yield in Embryonated Eggs

Kilbourne and Murphy (1960) were first to demonstrate that genetic reassortment could be used to derive viruses containing the antigenic properties of wild type viruses and the high yielding properties of laboratory strains. Schulman and Palese (1978) compared the growth properties of 47 reassortant viruses obtained from PR8 and HK viruses and found that the high-yielding property was associated with PR8 virus segments coding for NP and M1 and M2. When reassortment mixtures were inoculated into eggs at low dilution, there was enrichment for reassortant viruses containing PR8 virus segments coding for NP and M1, M2 but no selection for reassortants containing PR8 genes for P proteins or NS1, NS2. Baez *et al.* (1980) determined the genotypes of high-yielding reassortant viruses obtained after mixed infection with PR8 virus and different wild type viruses and found that the only RNA segment in common to all was RNA 7, coding for M1 and M2 proteins.

C. Sensitivity to Amantadine

Different strains of influenza A virus have been found to vary appreciably in their sensitivity to the inhibitory effects of amantadine (reviewed by Oxford, 1977). In two independent studies, involving analyses of genotype and amantadine sensitivity, it was found that the RNA segment coding for M1 and M2 proteins determines susceptibility or resistance (Lubeck *et al.*, 1978; Hay *et al.*, 1979). However, in another study, amantadine sensitivity was found to be determined either by the gene coding for HA or principally by the genes coding for NP and NA, depending on the assay system employed (Scholtissek and Faulkner, 1979).

VIII. Virulence in Man

Burnet and Bull (1943) predicted 40 years ago that the genetic instability of influenza viruses might be exploited to select for viruses attenuated for man by serial passage in an alien host, such as embryonated eggs. Nevertheless, the genetic basis for attenuation of influenza viruses in man remains a mystery (reviewed by Beare, 1982). Clearly, laboratory-derived *ts* and cold-adapted mutants have been shown to be less virulent for man than the strains from which they were derived (reviewed by Murphy *et al.*, 1980). However, these were all selected on the basis of their replication at different temperatures, and it is not clear whether their attenuation in man is attributable to specific mutations in particular genes or to the fact that they have an altered temperature-dependent phenotype which influences their replication at the high temperatures of the lower respiratory tract. In the latter case, other mutants bearing lesions in other genes but with the same temperature-dependent phenotypes might be expected to be equally attenuated. Though such viruses may be of great potential value in constructing suitable live virus vaccines, they do not as yet readily permit identification of particular genes among wild type viruses which confer virulence in man.

Attempts have also been made to identify genes associated with virulence in humans among reassortant viruses obtained from wild type viruses isolated from throat washings and attenuated laboratory strains, but the results of these trials are not easy to interpret, because of the necessarily small numbers of volunteers and of reassortants tested, as well as host variability in genetic background and immunologic experience with other influenza viruses. Consequently, an association of particular genes or combinations of genes with virulence is difficult to derive. Thus, X-31 virus, which derives all of its genes from PR8 virus (Baez *et al.*, 1980), a virus which is not infectious in man, was found to be only partly attenuated, whereas other H3N2 reassortant viruses deriving different combinations of PR8 virus genes were either fully attenuated or virulent (reviewed by Beare, 1982). Similarly, trials among volunteers with genotyped reassortants derived from H3N2 virus and an attenuated laboratory strain of H2N2 virus did not reveal any consistent genotype associated with either virulence or attenuation and reassortants of the same genotype were found to have detectable differences in virulence (reviewed by Beare, 1982).

It has been difficult as well to establish whether naturally occurring influenza A virus strains vary significantly in their inherent virulence for humans (reviewed by Kilbourne, 1980). Wright *et al.* (1980) obtained evidence from sequential epidemics of H3N2 and H1N1 viruses in non-immune young children which suggested that H3N2 virus infection was more frequently associated with febrile illness, but it remains to be seen whether similar observations are made in other outbreaks. On the basis of the predominance of H1N1 virus isolates containing RNA segments coding for P and NP proteins or P proteins, NP, M1 and M2 derived from co-circulating H3N2 viruses, Young and Palese (1979) suggested that these naturally occurring reassortants may have a selective advantage. However, it again remains to be seen whether these viruses continue to predominate.

Although attempts have been made for several decades to derive strains of human influenza viruses attenuated for man by serial passage in alien host cell systems, the results of these manipulations have proven to be variable (Beare, 1982). Moreover, even when deadaptation to the human host has been observed, it has not been feasible to define the genetic basis for the altered phenotype. Similarly, the mutational events required for adaptation to the human host of viruses resident in non-human populations have not been defined (reviewed by Kilbourne, 1980). Consequently, there is at present only extremely limited knowledge concerning specific mutations in particular genes affecting adaptation to particular hosts. In this light, the recent isolation and characterization of an H7N7 virus during a fatal disseminated influenza epizootic in seals is of great interest. Each of the 8 RNA segments of the virus was found by competition hybridization to most closely resemble the equivalent segment of other avian viruses. Nevertheless, the virus was found to replicate better in mammalian species than in avian species (Webster *et al.*, 1981). Although this altered host range could be due to a particular constellation of avian virus genes, it is also possible that adaptation was mediated by a stepwise series of mutational events.

Finally, the question of genetic determinants of transmissibility must be addressed. However, if little is known about the determinants of virulence, even

less is known about the genetic controls of transmissibility. There is no reason to assume that strains which are unusually pathogenic for man (if such exist) are also more likely to be transmitted. On the one hand, viruses capable of replicating to high titer would be expected to be at a selective advantage for transmission. On the other hand, transmissibility *per se* might be influenced by mutations on genes other than those which may play a role in affecting the severity of illness. In fact, evidence from a laboratory model of transmission of influenza virus infection in mice suggests that transmissibility can be distinguished from other attributes of virulence, such as peak pulmonary virus titers and extent of pneumonia (Schulman, 1967; 1968).

IX. Summary

1. Thus far, the most clearly defined viral determinant of virulence and host range is susceptibility of the viral hemagglutinin to host cell proteolytic activation of infectivity. Among avian viruses, this property can be related to the hemagglutinin structure, but in the case of WSN virus, viral neuraminidase facilities cleavage of HA by an unknown mechanism.

2. With respect to the contributions of other genes, results are far less clear and depend upon the particular parental viruses employed to obtain reassortant viruses.

3. Clonal variation in phenotype among reassortant viruses containing the same distribution of parental genes indicates that random mutational events may exert important influences.

4. In some systems, virulence appears to require effective interaction of polypeptides encoded by genes derived from a single parent.

5. Reassortant viruses obtained from two parental viruses which are avirulent in a particular host cell system may be more virulent than either parent. In contrast, reassortants obtained from one virulent and one avirulent parent are most often at least partially attenuated.

6. In most animal systems, virulence tends to be under polygenic control, which probably reflects selection for mutations on several segments during adaptation to a particular host.

7. The genetic determinants of virulence and transmissibility in man have not been defined.

References

Alexander, D. J., Allan, W. H., Parsons, D. G., Parsons, G.: The pathogenicity of four avian influenza viruses from fowls, turkeys and ducks. Res. Vet. Sci. *24,* 242–257 (1978).

Almond, J. W.: A single gene determines the host range of influenza virus. Nature (Lond.) *270,* 617–618 (1977).

Askonas, B. A., McMichael, A. J., Webster, R. G.: The immune response to influenza viruses and the problem of protection against infection. In: Basic and Applied Influenza Research (Beare, A. S., ed.), 157–188. Boca Raton: CRC Press 1982.

Baez, M., Palese, P., Kilbourne, E. D.: Gene composition of high-yielding influenza vaccine strains obtained by recombination. J. Infect. Dis. *141,* 362–365 (1980).

Bean, W. J., Webster, R. G.: Phenotype properties associated with influenza genome segments. In:

Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 685–692. New York: Academic Press 1978.

Bean, W. J., Sriram, G., Webster, R. G.: Electrophoretic analysis of iodine-labeled influenza virus RNA segments. Analyt. Biochem. *102,* 228–232 (1980).

Beare, A. S.: Research into the immunization of humans against influenza by means of living viruses. In: Basic and Applied Influenza Research (Beare, A. S., ed.), 211–234. Boca Raton: CRC Press 1982.

Bosch, F. X., Garten, W., Klenk, H.-D., Rott, R.: Proteolytic cleavage of influenza virus hemagglutinins: Primary structure of the connecting peptide between HA1 and HA2 determines proteolytic cleavability and pathogenicity of avian influenza viruses. Virology *113,* 725–735 (1981).

Burnet, F. M., Bull, D. R.: Changes in influenza viruses associated with adaptation to passage in chick embryos. Austral. J. Exp. Biol. Med. Sci. *21,* 55–69 (1943).

Burnet, F. M.: Genetic interactions between animal viruses. In: The Viruses (Burnet, F. M., Stanley, W. M., eds.), Vol. 3, 275–306. New York: Academic Press 1959.

Campbell, D., Sweet, C., Hay, A. J., Douglas, A., Skehel, J. J., Mason, T. J., Smith, H.: Genetic composition and virulence of influenza virus: Differences in facets of virulence in ferrets, between two pairs of recombinants with RNA segments of the same parental origin. J. Gen. Virol. *58,* 387–398 (1982).

Conti, G., Valcavi, P., Natali, A., Schito, G. C.: Different patterns of replication in influenza virus-infected KB cells. Arch. Virology *66,* 309–320 (1980).

Erickson, A. H., Kilbourne, E. D.: Mutation in the hemagglutinin of A/NWS/33 influenza virus recombinants influencing sensitivity to trypsin and antigenic reactivity. Virology *107,* 320–330 (1980).

Gething, M. J., White, J. M., Waterfield, M. D.: Purification of the fusion protein of Sendai virus: Analysis of the NH2-terminal sequences generated during precursor activation. Proc. Natl. Acad. Sci. U.S.A. 75, 2737-2740 (1978).

Haller, O.: A mouse hepatotropic variant of influenza virus. Arch. Virology *48,* 77–88 (1975).

Haller, O., Arnheiter, H., Gresser, I., Lindenmann, J.: Genetically determined interferon-dependent resistance to influenza in mice. J. Exp. Med. *149,* 601–612 (1979).

Hay, A. J., Kennedy, N. C. T., Skehel, J. J., Appleyard, G.: The matrix protein gene determines amantadine-sensitivity of influenza viruses. J. Gen. Virol. *42,* 189–191 (1979).

Hinshaw, V. S., Webster, R. G.: The natural history of influenza A viruses. In: Basic and Applied Influenza Research (Beare, A. S., ed.), 79–104. Boca Raton: CRC Press 1982.

Hrdy, D. B., Rubin, D. H., Fields, B. N.: Molecular basis of reovirus neurovirulence: Role of the M2 gene in avirulence. Proc. Natl. Acad. Sci. U.S.A. *79,* 1298–1302 (1982).

Huang, R. T. C., Wahn, K., Klenk, H.-D., Rott, R.: Fusion between cell membrane and liposomes containing the glycoproteins of influenza virus. Virology *104,* 294–302 (1980).

Huang, R. T. C., Rott, R., Klenk, H.-D.: Influenza viruses cause hemolysis and fusion of cells. Virology *110,* 243–247 (1981).

Israel, A.: Genotypic and phenotypic characterization of a mammalian cell-adapted mutant of fowl plague virus (FPV). J. Gen. Virol. *51,* 33–44 (1980).

Jennings, R., Potter, C. W., Teh, C. Z., Mahmud, M. I. A.: The replication of type A influenza viruses in the infant rat: A marker for virus attenuation. J. Gen. Virol. *49,* 343–354 (1980).

Kilbourne, E. D., Murphy, J. S.: Genetic studies of influenza viruses. I. Viral morphology and growth capacity as exchangeable genetic traits. Rapid in ovo adaption of early passage Asian strain isolates by combination with PR8. J. Exp. Med. *3,* 387–406 (1960).

Kilbourne, E. D.: Genetic dimorphism in influenza viruses. Characterization of stably assocuated hemagglutinin mutants differing in antigenicity and biological properties. Proc. Natl. Acad. Sci. U.S.A. *75,* 6258–6262 (1978).

Kilbourne, E. D.: Influenza: Viral determinants of the pathogenicity and epidemicity of an invariant disease of variable occurrence. Phil. Trans. R. Soc. (Lond.) *B 288,* 291–297 (1980).

Klenk, H.-D., Rott, R., Orlich, M., Blodorn, J.: Activation of influenza viruses by trypsin treatment. Virology *68,* 426–439 (1975).

Lazarowitz, S. G., Choppin, P. W.: Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. Virology *68,* 440–454 (1975).

Lenard, J., Miller, D. K.: PH dependent hemolysis by influenza virus, Semliki Forest virus, and Sendai virus. Virology *110,* 479–482 (1981).

Lubeck, M. D., Schulman, J. L., Palese, P.: Susceptibility of influenza A viruses to amantadine is influenced by the gene coding for M protein. J. Virol. *28,* 710–716 (1978).

Mayer, V., Schulman, J. L., Kilbourne, E. D.: Nonlinkage of neurovirulence exclusively to viral hemagglutinin in genetic recombinants of A/NWS (H0N1) influenza virus. J. Virol. *11,* 272–278 (1973).

Michaels, R. H., Mahmud, M. I. A., Coup, A. J., Jennings, R., Potter, C. W.: Influenza virus infection in newborn rats: A possible marker of attenuation for man. J. Med. Virol. *2,* 253–264 (1978).

Murphy, B. R., Markoff, L. J., Chanock, R. M., Spring, S. S., Maassab, H. F., Kendal, A. P., Cox, N. J., Levine, M. M., Douglas, R. G., Betts, R. F., Couch, R. B., Cate, T. R.: Genetic approaches to attenuation of influenza A viruses for man. Phil. Trans. R. Soc. (Lond.) *B 288,* 401–415 (1980).

Nakajima, S., Sugiura, A.: Neurovirulence of influenza virus in mice. II. Mechanism of virulence as studied in a neuroblastoma cell line. Virology *101,* 450–457 (1980).

Ogawa, T., Ueda, M.: Genes involved in the virulence of an avian influenza virus. Virology *113,* 304–313 (1981).

Oxford, J. S.: Specific inhibitors of influenza related to the molecular biology of virus replication. In: Chemoprophylaxis and Virus Infections of the Respiratory Tract (Oxford, J. S., ed.), Vol. 1, 140–187. Cleveland: CRC Press 1977.

Palese, P., Schulman, J. L.: Mapping of the influenza virus genome: Identification of the hemagglutinin and the neuraminidase genes. Proc. Natl. Acad. Sci. U.S.A. *73,* 2142–2146 (1976).

Ritchey, M. B., Palese, P., Schulman, J. L.: Differences in protein patterns of influenza A viruses. Virology *76,* 122–128 (1977).

Rott, R., Orlich, M., Scholtissek, C.: Correlation of pathogenicity and gene constellation of influenza A viruses. III. Non-pathogenic recombinants derived from highly pathogenic parent strains. J. Gen. Virol. *44,* 471–477 (1979).

Rott, R.: Genetic determinants for infectivity and pathogenicity of influenza viruses. Phil. Trans. R. Soc. (Lond.) *B 288,* 393–399 (1980).

Rott, R., Reinacher, M., Orlich, M., Klenk, H.-D.: Cleavability of hemagglutinin determines spread of avian influenza viruses in the chorioallantoic membrane of chicken embryo. Arch. Virol. *65,* 123–133 (1980).

Rott, R.: Determinants of influenza virus pathogenicity. Hoppe-Seyler's Z. Physiol. *363,* 1273–1282 (1982).

Rott, R., Orlich, M., Scholtissek, C.: Differences in the multiplication at elevated temperature of influenza virus recombinants pathogenic and nonpathogenic for chicken. Virology *120,* 215–224 (1982).

Scheid, A., Graves, M. C., Silver, S. M., Choppin, P. W.: Studies on the structure and function of paramyxovirus glycoproteins. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 181–193. New York: Academic Press 1978.

Scholtissek, C., Harms, E., Rohde, W., Orlich, M., Rott, R.: Correlation between RNA fragments of fowl plague virus and their corresponding gene functions. Virology *74,* 332–344 (1976).

Scholtissek, C., Rott, R., Orlich, M., Harms, E., Rohde, W.: Correlation of pathogenicity and gene constellation of an influenza A virus (fowl plague). I. Exchange of a single gene. Virology *85,* 74–80 (1977).

Scholtissek, C., Rohde, W., von Hoyningen, V., Rott, R.: On the origin of the human influenza virus subtypes H2N2 and H3N2. Virology *87,* 13–20 (1978).

Scholtissek, C., Faulkner, G.: Amantadine-resistant and amantadine-sensitive influenza A strains and recombinants. J. Gen. Virol. *44,* 807–815 (1979).

Scholtissek, C., Vallbracht, A., Flehmig, B., Rott, R.: Correlation of pathogenicity and gene constellation of influenza A viruses. II. Highly neurovirulent recombinants derived from non-neurovirulent or weakly neurovirulent parent virus strains. Virology *95,* 492–500 (1979).

Schulman, J. L.: Experimental transmission of influenza virus infection in mice. IV. Relationship of transmissibility of different strains of virus and recovery of airborne virus in the environment of infector mice. J. Exp. Med. *125,* 479–488 (1967).

Schulman, J. L.: The use of an animal model to study transmission of influenza virus infection. Am. J. of Public Health *58,* 2092–2096 (1968).

Schulman, J. L.: Immunology of influenza. In: The Influenza Viruses and Influenza (Kilbourne, E. D., ed.), 373–394. New York: Academic Press 1975.

Schulman, J. L., Palese, P.: Virulence factors of influenza viruses. WSN virus neuraminidase is required for productive infection of MDBK cells. J. Virol. *24,* 170–176 (1977).

Schulman, J. L., Palese, P.: Biological properties of influenza A/Hong Kong and PR8 viruses: Effects of genes for matrix protein and nucleoprotein on virus yield in embryonated eggs. In: Negative Strand Viruses and the Host Cell (Mahy, B. W. J., Barry, R. D., eds.), 663–674. New York: Academic Press 1978.

Skehel, J. J., Waterfield, M. D.: Studies on the primary structure of influenza virus hemagglutinin. Proc. Natl. Acad. Sci. U.S.A. *72,* 93–97 (1975).

Slemons, R. D., Easterday, B. C.: Host response difference among five avian species to an influenza virus A/turkey/Ontario 7732/66 (Hav5N?). Bull. W.H.O. *47,* 521–527 (1972).

Sugiura, A.: Influenza virus genetics. In: The Influenza Viruses and Influenza (Kilbourne, E. D., ed.), 171–213. New York: Academic Press 1975.

Sugiura, A., Ueda, M.: Neurovirulence of influenza virus in mice. I. Neurovirulence of recombinants between virulent and avirulent strains. Virology *101,* 440–449 (1980).

Teh, C., Jennings, R., Potter, C. W.: Influenza virus infection of newborn rats: Virulence of recombinant strains prepared from influenza virus strain A/Okuda/57. J. Med. Microbiol. *13,* 297–306 (1980).

Vallbracht, A., Scholtissek, C., Flehmig, B., Gerth, H.-J.: Recombination of influenza A strains with fowl plague virus can change pneumotropism for mice to a generalized infection with involvement of the central nervous system. Virology *107,* 452–460 (1980).

Webster, R. G., Hinshaw, V. S., Bean, W. J., van Wyke, K. L., Geraci, J. R., St. Aubin, D. J., Petursson, G.: Characterization of an influenza A virus from seals. Virology *113,* 712–724 (1981).

White, J., Kartenbeck, J., Helenius, A.: Membrane fusion activity of influenza virus. The EMBO J. *1,* 217–222 (1982).

Wright, P. F., Thompson, J., Karzon, D. T.: Differing virulence of H1N1 and H3N2 influenza strains. Amer. J. Epidemiol. *112,* 814–818 (1980).

Young, J. F., Palese, P.: Evolution of human influenza A viruses in nature: Recombination contributes to genetic variation of H1N1 strains. Proc. Natl. Acad. Sci. U.S.A. *76,* 6547–6651 (1979).

11

Molecular Epidemiology of Influenza Virus

P. Palese and *J. F. Young*

I. Introduction

Recent developments of new technologies have greatly advanced our ability to define the genetic constitution and molecular structure of viral agents, to determine similarities among strains and to identify differences in viral variants. Most important among these techniques are cloning and sequencing of nucleic acids, RNA fingerprinting, peptide mapping and antigenic analysis of proteins with monoclonal antibodies. The purpose of this review, the last in this book, is to discuss the epidemiology of influenza viruses in the context of these novel techniques. Special emphasis will be put on the evaluation of recent results and an examination of the epidemiological questions which may be answered using these molecular techniques.

II. Influenza A, B, and C Viruses and Associated Diseases

As discussed in previous chapters, influenza viruses can be classified into three types, A, B, and C. Strains of the same type share serologically cross-reactive M and NP proteins, which are the major components of the virus. Of the three types, the influenza A viruses appear to be the most important disease agents in man because of their association with pandemic influenza.

In the uncomplicated case, influenza involves infection of epithelial cells in the upper respiratory tract. Following a short incubation period of one to three days, the patient usually suffers a high fever spike with malaise, nausea, headache, chills and cough. Convalescence is lengthy but recovery in general is complete after two to four weeks. Accounts of this clinical picture of influenza A virus infections have been many and excellent reviews have been published (Douglas, 1975; Stuart-Harris and Schild, 1976; Noble, 1982).

The question then arises whether or not the clinical pattern described for influenza A virus infections resembles that caused by influenza B and C viruses. Clearly, there are differences among the three virus types with respect to their epidemiology, attack rates in the human population and, most importantly, the

ease with which the viruses can be isolated and identified (Kilbourne, 1975). However, Dykes *et al.* (1980) have recently reported a study of illness associated with influenza C viruses. Although the number of patients examined was small, these authors suggested that the disease was comparable to or even more severe than that caused by influenza A viruses. Along the same lines, recent reports suggest that the clinical manifestations of influenza B virus infections are not consistently different from those of influenza A viruses (van Voris *et al.*, 1982; Baine *et al.*, 1980; La Montagne, 1980). Consequently, it may be appropriate to question earlier assumptions (Joosting *et al.*, 1968; Davenport, 1976) that influenza B and C virus infections cause distinctly lower morbidity than influenza A virus infections.

Complications of influenza A virus infections include primary viral pneumonia and/or, on rare occasions, cardiac and neurological involvement. Recently, an additional severe complication of influenza virus infections has attracted much attention. Reye's syndrome may follow outbreaks of influenza A and B viruses. Halsey *et al.* estimated a frequency of 2.5 cases per 100,000 children infected with H1N1 influenza A viruses. An incidence rate ten times higher was estimated for children infected with influenza B viruses (Corey *et al.*, 1974). Although Reye's syndrome is well known for its hepatic and neurological abnormalities and its age-related incidence, little is known about its precise etiology; environmental factors and/or differences in the host response to infection seem to be more important than a specific phenotype of the infecting virus. In this context, it should be noted that in preliminary experiments influenza A and B viruses isolated from Reye's syndrome patients were indistinguishable from other strains when analyzed by SDS protein gel electrophoresis. (Krystal, Partin, and Palese, unpublished).

With the exception of work on Reye's syndrome, little has been done to establish the frequency and variety of complications following influenza B virus infections. Even less is known about possible complications following influenza C virus infections, although numerous studies have demonstrated antibody to the virus in the majority of adults (O'Callaghan *et al.*, 1980; Homma *et al.*, 1982).

III. Influenza in the 20th Century

For many years, influenza A viruses have received much attention from clinical and molecular epidemiologists (Kilbourne, 1973) because they are agents that can cause pandemics. These global epidemics are usually the results of the appearance of a new subtype strain. It is believed that there were at least three separate occasions in this century when viruses with novel HA/NA combinations took hold in the human population (Fig. 1). The first new antigenic combination appears to have arisen in the 1918 swine virus strain, which was responsible for a high incidence of morbidity and mortality. This "new" virus strain caused a pandemic which resulted in the death of an estimated 20 million people worldwide. Serological evidence suggests that the 1918 virus is related to viruses first isolated in 1931 from pigs.

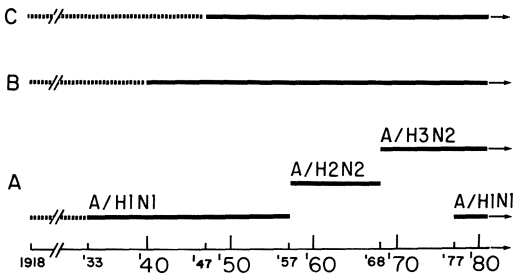


Fig. 1. Circulation of human influenza A, B, and C viruses since 1918, Broken lines indicate possible continued circulation of strains but no virus isolates are available from these periods. The subtype of an influenza A virus is defined by its surface proteins, the hemagglutinin *(H)* and the neuraminidase *(N)*. Three different A virus subtypes have been observed in this century. It should be noted that antigenic drift occurs among influenza B and C viruses as well as among strains belonging to the same influenza A subtype

Consequently, these strains are referred to as swine viruses or H1N1 viruses. It appears that viruses with hemagglutinins that were slightly different from the 1918 swine hemagglutinin continued to circulate for almost 40 years. In 1956, the H2N2 strains (Asian strains) emerged and replaced the H1N1 viruses. These viruses were not neutralized by antibodies directed against the earlier H1N1 strains. The H2N2 strains then circulated for eleven years until, in 1968, a new subtype strain appeared. The new H3N2 viruses (Hong Kong strains) retained the neuraminidase of the H2N2 strains and exhibited a previously unrecognized H3 hemagglutinin. Thus, from 1918 to 1968 three different subtype strains (H1N1, H2N2, and H3N2) circulated sequentially in the human population.

In addition to the appearance of these three novel influenza virus subtypes , an "old" influenza virus subtype emerged in 1977. These 1977 strains (Russian viruses) were unexpectedly found to possess surface proteins similar to H1N1 viruses which had circulated around 1950 (Kung *et al.*, 1978; Zhdanov *et al.*, 1978;, Kendal *et al.*, 1978). Since 1977 these "Russian" strains have coexisted in the population with the H3N2 viruses which had first appeared in 1968.

The epidemiology of influenza B and C viruses appears to be less complex (Fig. 1). Influenza B viruses have been isolated and characterized since 1940. During this 40 year period, there has been no indication that different influenza B virus subtypes exist. However, these analyses have shown that antigenic drift occurs in influenza B viruses in a manner similar to that of influenza A viruses (Chakraverty, 1971; Schild *et al.*, 1973). Most of the epidemiological studies involving influenza C viruses have been based on serologic analyses of different populations and less on the characterization of virus isolates. Primary virus isolation using standard laboratory procedures may have been less successful with influenza C viruses, since they require special growth conditions (Nerome and Ishida, 1978). From the available data it appears, however, that influenza C virus infections occur in an endemic pattern and not in the form of epidemic outbreaks. Although minor antigenic differences among influenza C viruses have been observed (Chakraverty, 1978), it is not clear whether these antigenic changes have any impact on the epidemiology of the virus.

IV. Surveillance of Influenza Viruses

An effective biochemical analysis of influenza virus field strains and the appropriate interpretation of the data are dependent on reliable epidemiological surveillance systems. International surveillance of human influenza has been under the aegis of the World Health Organization. These efforts are supported by national influenza virus centers throughout the world. In 1980, 101 national centers in 71 countries were collaborating with the two WHO influenza virus centers in Atlanta and London. In the USA, influenza viruses are collected by state or local health laboratories and an extensive nationwide surveillance system has developed. An excellent account of these ongoing activities is found in the Morbidity and Mortality Weekly Reports issued by the U.S. Public Health Service Centers for Disease Control in Atlanta.

Although few incidents have been published, our laboratories and others have received influenza virus samples which were later proven to be laboratory artifacts or laboratory contaminants; it is thus imperative that laboratories all over the world which are engaged in active surveillance and in the characterization of specimens adhere to good laboratory practices (WHO, Scientific Activities, 1981). Of particular importance is the establishment of competent laboratories, such as the Influenza Research Center in Houston, which have the appropriate technology and manpower to conduct reliable surveillance and which provide continuing information on the epidemiology of influenza viruses.

Methods of measuring the morbidity and mortality of influenza are based on a variety of indices. Morbidity estimates can be obtained using absenteeism values, reportings of general physicians insurance claims and records, or on the number of influenza virus isolates obtained from a population. Mortality figures may be calculated based on the excess mortality in a given month or on data directly extracted from death certificates. Thus, many methods can be used for measuring the impact of influenza on the population (Noble, 1982).

V. The 1977 H1N1 Viruses

In the following section it will be demonstrated how biochemical techniques have successfully been used to further analyze an intriguing epidemiological event. The example cited concerns the reappearance of the H1N1 influenza viruses in 1977.

The assumption that influenza A viruses of only one subtype circulate at a particular time was challenged when, starting in 1977, H1N1 as well as H3N2 strains were isolated from man. A second unique feature of this event became clear when it was found that H1N1-related disease was virtually absent among adults. Although sporadic illnesses among adults born before 1957 have been reported, immunity resulting from previous H1N1 infections apparently prevented infection and/or illness among adults. Detailed serologic analyses revealed that both the HA as well as the NA of the 1977 strains were similar to those of 1950 strains. Specifically, ferret sera appeared to provide excellent probes in hemagglutination inhibition tests

(Kendal *et al.*, 1978) and monoclonal antibodies were used to demonstrate that the antigenic determinants of the 1977 strains were homologous to those of 1950 strains (Webster *et al.*, 1979). Thus it was not surprising that previous exposure to the antigen provided protection against the 1977 H1N1 strains.

Further biochemical and genetic analysis of the 1977 strains introduced a new dimension into the complex pattern of influenza epidemiology. At that time it was widely held that antigenic drift was the consequence of point mutations and immune selection and that any radically new subtype resulted from recombination between human and animal viruses (Laver and Webster, 1975). Consequently it was unexpected when oligonucleotide map analysis of the RNAs of the 1977 strains revealed that all eight genes were practically identical to those of the earlier human strains (Nakajima *et al.*, 1978). This finding was confirmed using RNA-RNA hybridization techniques (Scholtissek *et al.*, 1978 b; Bean *et al.*, 1980) and suggested that the 1977 strains were not recombinants.

This latter point was questioned by Kozlov *et al.* (1981). These authors suggested, based on oligonucleotide mapping data, that the 1977 strains were recombinants between 1947 and 1950 strains. Since the oligonucleotide mapping technique used by Kozlov *et al.*, was hardly sensitive enough to differentiate between the nearly identical genes of a 1947 and a 1950 strain it is difficult to accept the interpretation of these authors. Rather, the data by Kozlov *et al.*, merely go along with the evidence presented by other laboratories that the 1977 strains closely resemble the earlier H1H1 strains. Clearly, the presently available data do not permit identification of the precise precursor variant which gave rise to the 1977 strains, but the data do show that a strain circulating around 1950 (plus or minus several years) was the direct forerunner of these strains.

The data also suggest that old influenza viruses previously in circulation may be "reactivated" through natural pathways or through the accidental introduction of laboratory strains. Whatever the precise circumstances of the reemergence of the 1977 strains are, it is now clear from these biochemical studies that a change in the epidemiologic pattern from one influenza A virus subtype to another need not necessarily involve recombination (reassortment) between different (human and animal) influenza virus strains.

VI. Molecular Epidemiology of Influenza in Animals

Influenza A viruses have been isolated from a variety of animals such as pigs, horses, seals, and domestic and wild birds. Virus strains belonging to thirteen HA and nine NA subtypes have been identified and many different HA/NA combinations have been observed (Hinshaw and Webster, 1982). In particular, extensive surveillance programs in bird populations resulted in the isolation of a plethora of avian strains and there is little doubt now that birds represent the largest natural reservoir of influenca A viruses. Ongoing studies promise to unravel the significance of this animal reservoir with respect to the formation of new strains causing disease in man. In essence, the identical techniques which were applied to the study of human influenza viruses can be used in the characterization of animal strains.

Specifically, oligonucleotide mapping of the RNAs (Desselberger *et al.*, 1978), sequencing of RNAs and cloned genes (Air *et al.*, 1981; Fang *et al.*, 1981), peptide mapping, protein and RNA gels (Laver and Webster, 1973; Palese and Ritchey, 1977; Hinshaw *et al.*, 1980) and serologic techniques (Schild *et al.*, 1980; Shortridge, 1982) have been successfully used to differentiate animal influenza viruses. However, the animal influenza viruses have been less well studied than the strains obtained from man. This includes structural data on the viruses as well as data on epidemiological characteristics of animal strains and the immune and pathogenic responses of the infected animal host. Future studies will undoubtedly be directed at answering some of these questions. It will also be important to compare the variation rate of animal strains in horses, pigs and birds with that of human strains and to define the factors contributing to the pathogenicity of animal strains. Finally, efforts will have to be made to understand the mechanisms influencing host species specificity among influenza A viruses and to identify determinants that govern the interactions of animal and human virus strains.

It is commonly accepted that influenza B and C viruses are largely restricted to man. However, there have been occasional reports (Kawano *et al.*, 1978) suggesting the presence of antibodies to influenza B viruses in animals such as horses, swine or dogs. The frequency of measurable immune responses has been low, consistent with a limited incidence of influenza B virus infections in animals. Similarly, a few cattle antisera were identified which contained anti-influenza C virus antibodies (Kawano *et al.*, 1978). More significantly, however, a recent report describes the isolation of influenza C viruses from pigs in China (Guo *et al.*, 1982). This finding may indeed indicate that animals can provide a reservoir for influenza C viruses. It will be important to see whether improved isolation techniques can result in more frequent recovery of influenza C viruses from animals.

VII. Mechanisms Contributing to Variation in Influenza Virus Field Strains

Tables 1 and 2 summarize the available information on the genetic mechanisms contributing to variation among influenza A, B and C viruses. These data are based on the work of many laboratories and include results derived by many techniques.

A. Point Mutations

Variation in the HA genes of the influenza A viruses has been extensively studied. The HA genes of the three human subtype viruses, H1, H2 and H3, as well as those of two avian strains have been sequenced to date. A remarkable collection of sequencing data is available for the HAs of H3 field strains (chapter 5). These results have helped to delineate the antigenic sites in the three-dimensional structure of the HA and to define the sequential nature of mutations in the HAs of the predominant field isolates (Wilson *et al.*, 1981; Wiley *et al.*, 1981).

Table 1. *Mechanisms responsible for genetic variability in influenza A viruses*

Variation within subtype strains
1. Point mutations in genes coding for surface as well as nonsurface proteins.
2. Short deletions/insertions in genes coding for hemagglutinins and neuraminidases.
3. Reassortment (shuffling) of genes coding for nonsurface proteins.
Variation leading to different subtypes
1. Reassortment of surface protein genes.
2. Reemergence of previously circulating strains.

In contrast to influenza viruses, other viruses such as herpes viruses or rhinoviruses do not have a single variant which is prevalent at one particular time, rather, many variants cocirculate at the same time. Although influenza viruses undergo sequential changes, this does not exclude the simultaneous presence of more than one variant in the population at any one time. Over a period of several years, however, lineages develop suggesting that the predominant variants arise by sequential changes with time. An extensive analysis of H3 hemagglutinins using sequencing techniques greatly enhanced our understanding of the evolution of influenza virus hemegglutinins during a single subtype period (Both and Sleigh, 1980, 1981; see Fig. 2 and chapter 5).

The fact that not only the HAs but also nonsurface proteins undergo variation can be seen by protein analysis of different field variants on SDS polyacrylamide gels (Fig. 3). Clearly, corresponding proteins exhibit different electrophoretic migration patterns in this system. A more definitive analysis of variation can be made using recombinant DNA techniques. Following cloning and sequencing, an evolutionary tree was determined for the NS genes of several field strains (Fig. 4). The results indicate that the NS genes, like the HA genes, also evolve along a common lineage. A rate of variation of 2.2 to 3.4% sequence divergence per ten years in the NS genes can be calculated by using the A/PR/8/34, A/FM/147, A/FW/1/50 and A/Udorn/72 NS gene sequences (Krystal *et al.*, 1983 b). Obviously, this value represents only an estimate and ignores sampling problems which are inherent in such an analysis. The variation rate observed among the NS genes can be compared with rates calculated for the HA genes. The rate of 4.5 to 6.5% per 10 years for the latter genes appears to be higher than that observed for the NS genes. Since immune response is probably greater to the HAs than to the NS proteins, the higher overall variation rate in the HA genes may in part be the result of the host selection process. An alternative view is that more sequence data, including those for other

Table 2. *Variation among influenza B virus strains*

1. Point mutations.
2. Short deletions/insertions in hemagglutinin genes.
Variation among influenza C virus strains
Not sufficient data available

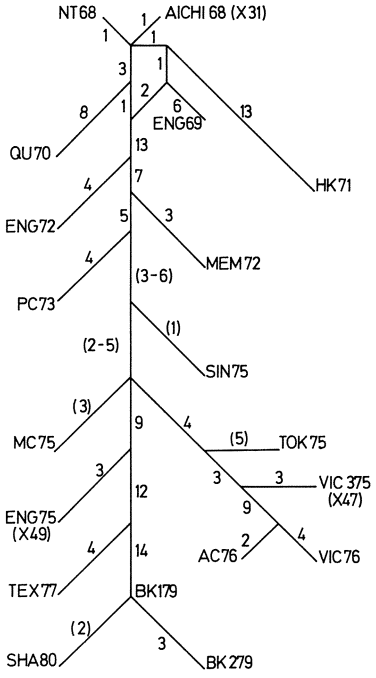


Fig. 2. Interrelationship of H3 subtype strains as determined from minimum mutational distances among their HA1 coding regions. Numbers signify the HA1 base differences between strains; those in brackets are obtained from partial sequences only. (From Both, Sleigh, Cox, and Kendal, submitted)

genes, are required before average variation rates for influenza virus RNAs can be reliably determined. A recent sequence analysis of three influenza B virus HAs has allowed for an examination of variation in type B viruses (Krystal *et al.*, 1983 a). As seen with A virus HAs, variation proceeds mostly through the sequential accumulation of amino acid substitutions within the HA1 portion of the HA molecule (Fig. 5). However, the rate of variation observed among the B/Lee/40, B/Md/59 and B/HK/8/73 HAs appears to be significantly slower than the observed rate of variation among influenza A virus HAs. The overall amino acid change in the HA 1s of the influenza B viruses is 2% per 10 years while the HA1s of H3 influenza A viruses vary by 9.2% per 10 years (Both and Sleigh, 1980). Although the sequence of the HA1 of the influenza C virus C/Cal/76 has been obtained recently (Creager *et al.*, unpublished), little can be said about variation among the HAs of influenza C viruses. Comparison of this sequence with the partial HA sequences of C/Taylor/47 virus shows the occurrence of single point mutations, but a more detailed analysis must await more experimental data.

B. Recombination (Reassortment)

Many data suggest that new subtype strains of influenza A viruses emerge through recombination (reassortment) of genes among different strains. Specifically, trans-

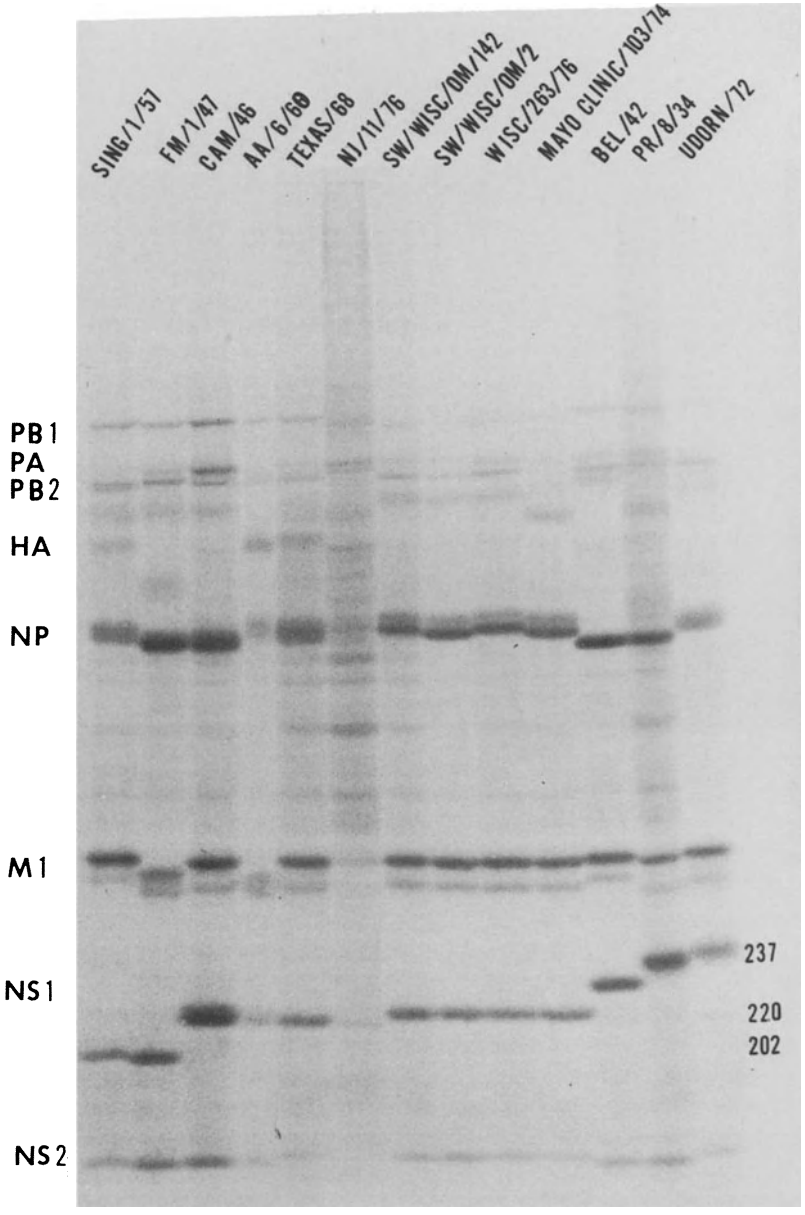


Fig. 3. Comparison of influenza virus proteins from infected MDCK cells. The 35S-methionine labeled cell extracts were separated in 5 to 13% polyacrylamide gradient gels containing sodium dodecylsulfate (Young and Palese, 1979). The polypeptide products coded for by the different viral genes are identified on the left of the autoradiogram. The numbers on the right indicate the varying amino acid lengths of different NS1 polypeptides. Migration differences of homologous polypeptides reflect differences in size as well as in secondary structure (Krystal *et al.*, 1983 b)

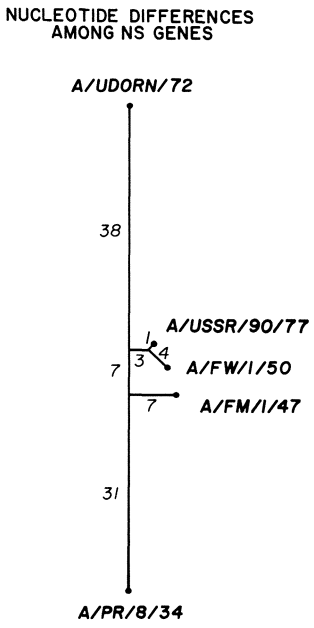


Fig. 4. Evolutionary tree based on nucleotide differences among NS genes. Five influenza virus strains isolated in 1934, 1947, 1950, 1972 and 1977, respectively, were used for the analysis. The 890 nucleotide long NS genes of the five strains were completely sequenced and compared with each other. Numbers indicate nucleotide changes. It appears that mutations in the NS genes are sequential, with the exception of the A/USSR/90/77 strain. (This confirms the earlier finding that the A/USSR/90/77 virus is closely related to strains circulating around 1950.) The rate of change is approximately two mutations per year. The sequences of the entire genes (approximately 880 nucleotides long) were used for the analysis

Fig. 5. Evolutionary tree of influenza B viruses using nucleotide differences among the HA genes. The numbers indicate nucleotide differences among the three HA genes of B/Lee/40, B/Md/59 and B/HK/8/73 viruses. The sequences of the entire genes (approximately 1880 nucleotides long) were used for the analysis

fer of the H3 hemagglutinin from an animal virus to the earlier H2N2 strains by recombination may have generated the new human subtype strain H3N2 (Laver and Webster, 1972; Laver, 1978; Scholtissek *et al.*, 1978; Fang *et al.*, 1981). Furthermore, it is likely that the H2N2 strains emerged from H1N1 strains through a recombinational event since the former contain genes coding for the nonsurface proteins derived from the H1N1 viruses (Scholtissek *et al.*, 1978 a). Among avian strains, recombination (reassortment) most likely also plays a role in the generation of new viruses (Desselberger *et al.*, 1980; Hinshaw *et al.*, 1980) and there is little doubt that similar recombinational events may occur in other hosts.

In addition to the exchange of the surface protein genes, which leads to antigenic shift (Table 3), cocirculating strains may exchange genes coding for the nonsurface proteins. The resulting recombinant (reassortant) retains the HA/NA subtype of one of the parents. For example, oligonucleotide and peptide maps of the genes and gene products, respectively, of the A/Cal/10/78 (H1N1) virus isolated in Los Angeles in December, 1978, suggested that only the HA, NA, M and NS genes of this virus are derived from the earlier H1N1 strains and that the four genes coding for the core complex (PB1, PB2, PA and NP) came from an H3N2 parent. This finding was unexpected, since viruses of the same subtype were previously thought to be closely related through point mutations only. Strains similar to the A/Cal/10/78 prototype have been found to circulate widely throughout the world (Young and Palese, 1979; Bean *et al.*, 1980). In addition to the A/Cal/10/78 (H1N1) recombinant type, H1N1 strains were also identified which possess five "H3N2" genes and only the HA, NA, and the NS genes from an A/USSR/90/77-like strains. The prototype strain for recombinants (reassortants) of this type ist the A/Aberdeen/v1340/78 (H1N1) strain (Fig. 6). The finding of novel recombinants with genotypes different from that of the A/Cal/10/78 virus suggests that reassortment may occur frequently among co-circulating strains and that, in addition to mutation, reassortment contributes to genetic variation and possibly to biological diffeences in human influenza viruses belonging to a single subtype. Since cocirculating strains can be genetically very similar, it may be difficult to assess the impact recombination

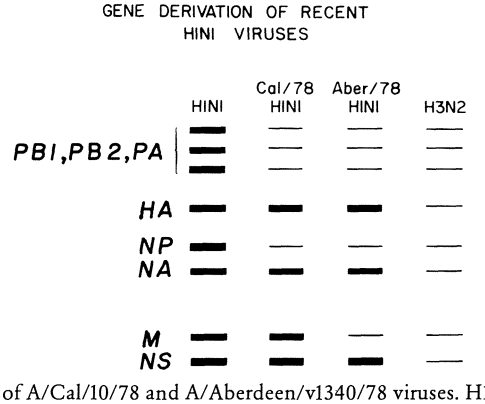


Fig. 6. Gene derivation of A/Cal/10/78 and A/Aberdeen/v1340/78 viruses. H1N1 genes of A/USSR/90/77-like strains are depicted with heavy bars. Genes similar to those of the H3N2 strain A/Tex/1/77 are noted with narrow lines. The Cal/78 and the Aber/78 strains derive four and five genes, respectively, from the H3N2 parents (from Young and Palese, 1979 and unpublished)

Table 3. *Sequence homologies of HAs from different influenza viruses*

Comparison[a]	Amino acid conservation, %[b]	
	HA1	HA2
H1 vs. H2	58	79
H1 vs. H3	35	53
H2 vs. H3	36	50
B vs. H1	24	39

[a] H1, A/PR/8/34; H2, A/Jap/305/57; H3, A/Aichi/2/68; and B, B/Lee/40. The sequences of the H1, H2, H3, and B virus HAs are from Winter *et al.*, 1981; Gething *et al.*, 1980; Verhoeyen *et al.*, 1980; Krystal *et al.*, 1982.

[b] Homologies were calculated as numbers of homologous residues/total number of residues. The total number of residues is the average of the length of the two sequences. The signal peptide sequences are not used for this comparison.

(reassortment) may have during the cocirculation of epidemic strains. It is conceivable, however, that this mechanism explains in part the accelerated selection of epidemiologically significant variants of influenza virus compared with other viruses.

C. Deletions/Insertions

Comparison of nucleotide sequences of different H3 HA genes demonstrated that alignment of invariant amino acids required deleting or inserting one or more triplets. For example, Vic/3/75 HA contains (at amino acid position 8) an additional asparagine which is not present in any other H3 HA sequenced so far (Verhoeyen *et al.*, 1980). It is not known whether the apparent insertion in the Vic/3/75 HA changes the antigenic character of the molecule but it is likely that in another instance a short deletion in an HA induces such an antigenic change. The B/HK/8/73 HA lacks two amino acids at positions 158 and 159 which are present in several other B virus HAs. Since a structural analysis suggests that amino acid positions 158 and 159 define an antigenic loop in the HA, it is probable that this deletion is accompanied by an antigenic change in the molecule (Krystal *et al.*, 1982, 1983 a). Thus, short deletions and insertions appear to contribute to the genetic variation of closely related HAs.

More extensive deletions/insertions are observed when the HAs of different type or subtype strains are compared (Hiti *et al.*, 1981; Krystal *et al.*, 1982). Since it is assumed that all influenza virus HAs, including those of the B and C viruses, derive from a common ancestor, additions or deletions of nucleotides appear to be common events contributing to the evolution of influenza virus HAs. In this respect it should be mentioned that the variation observed in the HAs of influenza viruses, which have such a narrow host range in mammalian and avian cells, is as great as that seen in certain proteins which are found in vastly different eucaryotic species. For example, the amino acid sequences of cytochrome C molecules from over 60 different eucaryotic species have been determined and only 27 positions in the chain of approximately 100 amino acids were found to be identical in all species examined (Lehninger, 1982). The different species analyzed include plants, yeasts,

amphibians, birds and mammals. It is thus quite remarkable that such a striking evolutionary divergence has been observed among HA genes of influenza virus strains isolated from man alone.

Deletions/insertions have also been observed in the NA genes. Partial sequencing of the first 200 to 300 3′-terminal sequences of different NI genes revealed length variation in this region (Blok and Air, 1982 b) and comparison of the completed sequences of N1 and N2 genes also revealed deletions/insertions during the evolution of these genes (Fields *et al.*, 1980; Hiti and Nayak, 1982; Markoff and Lai, 1982).

As yet there is no evidence that deletions/insertions play a role in the variation of nonsurface protein genes of recently isolated influenza A viruses. However, such events must have taken place during the evolution of influenza A, B, and C viruses from a common ancestor. For example, corresponding M and NS genes of influenza A and B viruses have been shown to contain homologous sequences but to vary in overall size (Desselberger and Palese, 1978; Briedis and Lamb, 1982; Briedis *et al.*, 1982).

D. Recycling of Genes

Another interesting aspect of variation in influenza virus genes is the evidence for possible antigenic recycling. Serum samples collected from aged persons before the H3N2 pandemic in 1968 contained antibodies to H3-like agents and stored serum samples from people born before 1887 cross-reacted with H2N2 strains isolated after 1957 (Davenport *et al.*, 1969; Masurel, 1969). These data, supported by the reemergence of the 1977 H1N1 viruses discussed above, may indicate that there is only a limited number of influenza virus HA and NA genes which may be found in human viruses and which contribute to variation of these strains.

VIII. Outlook for the Future

The influenza viruses represent a family within the viral kingdom which show a unique epidemiological pattern in nature. Great strides have been made in recent years in elucidating the structure, replication and evolution of these viruses.

Biochemical techniques are now being used as epidemiological tools. These procedures have helped in understanding the mechanisms of antigenic change in the hemagglutinins and the neuraminidase of the viruses, as well as the genetic changes in the genes coding for nonsurface proteins. Molecular techniques were also helpful in defining the genetic makeup of the H1N1 viruses which reappeared in 1977, and the H1N1 recombinant (reassortant) strains which circulated in subsequent years. In the latter instance it was shown that reassortment between cocirculating human strains can contribute to genetic variation. Analysis of the gene derivation of the pandemic H2N2 and H3N2 viruses has demonstrated that some of the influenza virus genes are retained in all human strains even though new genes have been introduced through recombination (reassortment). This may point to a role for these genes in determining host specificity. Although we have

learned a great deal about influenza viruses through the application of these techniques, continued interaction between clinical epidemiologists and molecular biologists will be needed in order to understand the complex behavior of this organism.

References

Air, G. M.: Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus. P.N.A.S. *78,* 7639–7643 (1981).

Baine, W. B., Luby, J. P., Martin, S. M.: Severe illness with influenza B. Am. J. Med. *68,* 181–189 (1980).

Bean, W. J, jr., Cox, N. J., Kendal, A. P.: Recombination of human influenza A viruses in nature. Nature *284,* 638–640 (1980).

Blok, J., Air, G. M.: Block deletions in the neuraminidase genes from some influenza A viruses of the N1 subtype. Virology *118,* 229–234 (1982).

Both, G. W., Sleigh, M. J.: Complete nucleotide sequence of the hemagglutinin gene from a human influenza virus of the Hong Kong subtype. Nucl. Acids Res. *8,* 2561–2575 (1980).

Both, G. W., Sleigh, M. J.: Conservation and variation in the hemagglutinin of Hong Kong subtype influenza viruses during antigenic drift. J. Virol. *39,* 663–672 (1981).

Briedis, D. J., Lamb, R. A.: The influenza B virus genome sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS1 and NS2 proteins. J. Virol. *42,* 186–193 (1982).

Briedis, D. J., Lamb, R. A., Choppin, P. W.: Sequence of RNA segment 7 of the influenza B virus genome: Partial amino acid homology between the membrane proteins (M1) of influenza A and B viruses and conservation of a second open reading frame. Virology *116,* 581–588 (1982).

Chakraverty, P.: Antigenic relationship between influenza B viruses. Bull. W.H.O. *45,* 755–766 (1971).

Chakraverty, P.: Antigenic relationship between influenza C viruses. Arch. Virol. *58,* 341–348 (1978).

Corey, L., Rubin, R. J., Hattwick, M. A. W., Noble, G. R., Cassidy, E.: A nationwide outbreak of Reye's syndrome: its epidemiologic relationship to influenza B. Am. J. Med. *61,* 615–625 (1976).

Davenport, F. M., Minuse, E., Hennessy, A. V., Francis, T.: Interpretations of influenza antibody patterns of man. Bull. W.H.O. *41,* 453–460 (1969).

Davenport, F. M.: Influenza viruses. In: Viral Infection of Humans (Evans, A. S., ed.), 273–296. New York: Plenum Medical Book & Co. 1976.

Desselberger, U., Nakajima, K., Alfino, P., Pedersen, F.-S., Haseltine, W., Hannoun, C., Palese, P.: Biochemical evidence that "new" influenza virus strains in nature may arise by recombination (reassortment). P.N.A.S. *75,* 3341–3345 (1978).

Desselberger, U., Palese, P.: Molecular weights of RNA segments of influenza A and B viruses. Virology *88,* 394–399 (1978).

Douglas, R. G.: Influenza in man. In: The Influenza Viruses and Influenza (Kilbourne, E. D., ed.), 395–447. New York: Academic Press 1975.

Dykes, A. C., Cherry, J. D., Nolan, C. E.: A clinical, epidemiologic, serologic and virologic study of influenza C virus infection. Arch. Intern. Med. *140,* 1295–1298 (1980).

Fang, R., Min Jou, W., Huylebroeck, D., Devos, R., Fiers, W.: Complete structure of A/duck/Ukraine/63 influenza hemagglutinin gene: Animal virus as progenitor of human H3 Hong Kong 1968 influenza hemagglutinin. Cell *25,* 315–323 (1981).

Fields, S., Winter, G., Brownlee, G. G.: Structure of the neuraminidase gene in human influenza virus A/PR/8/34. Nature *290,* 213–217 (1981).

Gething, M. J., Bye, J., Skehel, J. J., Waterfield, M.: Cloning and DNA sequence of double stranded copies of hemagglutinin genes from H2 and H3 strains elucidate antigenic shift and drift in human influenza virus. Nature *287,* 301–306 (1980).

Guo, Y., Jin, F., Wang, M., Wang, P., Zhu, J.: Influenza C virus isolated from pigs in China. Kexue Tongbao *27,* 1118–1121 (1982).

Halsey, N. A., Hurwitz, E. S., Meiklejohn, G., Todd, W., Todd, J. K., Edell, T., Graves, P., McIntosh, K.: Epidemic Reye's syndrome associated with influenza A (H1N1) infections. In: Current Chemo-

therapy and Infectious Diseases (Nelson, J. D., Grassi, C., eds.), Vol. 2, 1207. Washington, D. C.: American Society for Microbiology 1980.
Hinshaw, V. S., Bean, W. J., Webster, R. G., Sriram, G.: Genetic reassortment of influenza A viruses in the intestinal tract of ducks. Virology *102,* 412–419 (1980).
Hinshaw, V. S., Webster, R. G.: The natural history of influenza A viruses. In: Basic and Applied Influenza Research (Beare, A. S., ed.), 79–104. West Palm Beach, Fl.: CRC Press 1982.
Hiti, A. R., David, A. R., Nayak, D. P.: Complete sequence analysis shows that the hemagglutinins of the H0 and H2 subtypes of human influenza virus are closely related. Virology *111,* 113–124 (1981).
Hiti, A. L., Nayak, D. P.: Complete nucleotide sequence of the neuraminidase gene of influenza virus A/WSN/33. J. Virol. *41,* 730–734 (1982).
Homma, M., Ohyama, S., Katagiri, S.: Age distribution of the antibody to type C influenza virus. Microbiol. Immunol. *26,* 639–642 (1982).
Joosting A. C. C., Head, B., Bynde, M. L., Tyrrell, D. A. J.: Production of common colds in human volunteers by influenza C virus. Br. Med. J. *4,* 153–154 (1968).
Kawano, J., Onta, T., Kida, H., Yanagawa, R.: Distribution of antibodies in animals against influenza B and C viruses. Jap. J. Vet. Res. *26,* 74–80 (1978).
Kendal, A. P., Noble, G. R., Skehel, J. J., Dowdle, W. R.: Antigenic similarity of influenza A (H1N1) viruses from epidemics in 1977–1978 to "Scandinavian" strains isolated in epidemics of 1950–1951. Virology *89,* 632–636 (1978).
Kendal, A. P., Joseph, J. M., Kobayashi, G., Nelson, D., Reyes, C. R., Ross, M. R., Sarandria, J. L., White, R., Woodall, D. F., Noble, G. R., Dowdle, W. R.: Laboratory based surveillance of influenza virus in the United States during the winter of 1977–1978. Am. J. Epidemiol. *110,* 449–461 (1979).
Kilbourne, E. D.: The molecular epidemiology of influenza. J. Infect. Dis. *127,* 478–487 (1973).
Kilbourne, E. D. (ed.): The Influenza Viruses and Influenza. New York: Academic Press 1975.
Kozlov, J. V., Gorbulev, V. G., Kurmanova, A. G., Bayev, A. A., Shilov, A. A., Zhdanov, V. M.: On the origin of the H1N1 (A/USSR/90/7 influenza virus. J. Gen. Virol. *56,* 437–440 (1981).
Krystal, M., Elliott, R. M., Benz, E. W., Young, J. F., Palese, P.: Evolution of influenza A and B viruses: Conservation of structural features in the hemagglutinin genes. P.N.A.S. *79,* 4800–4804 (1982).
Krystal, M., Young, J. F., Palese, P., Wilson, I. A., Skehel, J. J., Wiley, D. C.: Sequential mutations in the hemagglutinins of influenza B virus isolates: Definition of antigenic domains. P.N.A.S. *80,* 4527–4531 (1983 a).
Krystal, M., Buonagurio, D., Young, J. F., Palese, P.: Sequential mutations in the NS genes of influenza virus field strains. J. Virol. *45,* 547–554 (1983 b).
Kung, H. C., Jen, K. F., Yuan, W. C., Tien, S. F., Chu, C. M.: Influenza in China in 1977: Recurrence of influenza virus A subtype H1N1. Bull. W.H.O. *56,* 913–918 (1978).
La Montagne, J. R.: Summary of a workshop on influenza B viruses and Reye's syndrome. J. Infect. Dis. *142,* 452–465 (1980).
Laver, W. G.: Crystallization and peptide maps of neuraminidase "heads" from H2N2 and H3N2 influenza virus strains. Virology *86,* 78–87 (1978).
Laver, W. G., Webster, R. G.: Studies on the origin of pandemic influenza. Virology *48,* 445–455 (1972).
Lehninger, A. L.: Principles of Biochemistry. New York: Worth Publishers 1982.
Markoff, L., Lai, C.-J.: Sequence of the influenza A/Udorn/72 (H3N2) virus neuraminidase gene as determined from cloned full length DNA. Virology *119,* 288–297 (1982).
Masurel, N.: Relation between Hong Kong virus and former human A2 isolates and the A/equi2 virus in human sera collected before 1957, Lancet *I,* 907–910 (1969).
Nakajima, K., Desselberger, U., Palese, P.: Recent human influenza A viruses are closely related genetically to strains isolated in 1950. Nature *274,* 334–339 (1978).
Nerome, K., Ishida, M.: The multiplication of an influenza C virus in an established line of canine kidney (MDCK) cells. J. Gen. Virol. *39,* 179–181 (1978).
Noble, G. R.: Epidemiological and clinical aspects of influenza. In: Basic and Applied Influenza Research (Beare, A. S., ed.), 11–50. West Palm Beach, Fl.: CRC Press 1982.
O'Callaghan, R. J., Gohd, R. S., Labat, D. D.: Human antibody to influenza C virus: Its age-related distribution and distinction from receptor analogs. Infect. and Immun. *30,* 500–505 (1980).
Palese, P., Ritchey, M. B.: Polyacrylamide gel electrophoresis of the RNAs of new influenza virus

strains: An epidemiological tool. In: Proceedings of the International Symposium on Influenza Immunization II (Develop. Biol. Standard., Vol. 39, 411–415). Basel: S. Karger 1977.
Palese, P., Young, J. F.: Variation of influenza A, B, and C viruses. Science *215,* 1468–1474 (1982).
Schild, G. C., Pereira, M. S., Chakraverty, P., Coleman, M. T., Dowdle, W. R., Chang, W. K.: Antigenic variants of influenza B virus. Br. Med. J. *4,* 127–131 (1973).
Schild, G. C., Newman, R. W., Webster, R. G., Major, D., Hinshaw, V. S.: Antigenic analysis of influenza A virus surface antigens: Considerations for the nomanclature of influenza virus. Arch. Virol. *63,* 171–184 (1980).
Scholtissek, C., Rohde, W., von Hoyningen, V., Rott, R.: On the origin of the human influenza virus subtypes H2N2 and H3N2. Virology *87,* 13–20 (1978 a).
Scholtissek, C., von Hoyningen, V., Rott, R.: Genetic relatedness between the new 1977 epidemic strains (H1N1) of influenza and human influenza strains isolated between 1947 and 1957 (H1N1). Virology *89,* 613–617 (1978 b).
Shortridge, K. F.: Avian influenza A viruses of southern China and Hong Kong: Ecological aspects and implications for man. Bull. W. H. O. *60,* 129–135 (1982).
Stuart-Harris, G., Schild, C.: Influenza: The Viruses and the Disease. London: Edward Arnold 1982.
Van Voris, L. P., Belshe, R. B., Shaffer, J. L.: Nosocomial influenza B virus infection in the elderly. Ann. intern. Med. *96,* 153–158 (1982).
Verhoeyen, M., Fang, R., Min Jou, W., Devos, R., Huylebroeck, D., Saman, E., Fiers, W.: Antigenic drift between the hemagglutinin of the Hong Kong influenza strains A/Aichi/2/68 and 75. Nature *286,* 771–776 (1980).
Webster, R. G., Laver, W. G.: Antigenic variation of influenza viruses. In: The Influenza Viruses and Influenza (Kilbourne, E. D., ed.), 269–314. New York: Academic Press 1975.
Webster, R. G., Kendal, A. P., Gerhard, W.: Analysis of antigenic drift in recently isolated influenza A (H1N1) viruses using monoclonal antibody preparations. Virology *96,* 258–264 (1979).
W.H.O. Scientific Activities. Evaluating the authenticity of unusual influenza virus isolates. Bull. W.H.O. *59,* 846–847 (1981).
Wiley, D. C., Wilson, I. A., Skehel, J. J.: Structural identification of the antibody-binding sites of Hong Kong influenza hemagglutinins and their involvement in antigenic variation. Nature *289,* 373–378 (1981).
Wilson, I. A., Skehel, J. J., Wiley, D. C.: Structure of the hemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. Nature *289,* 366–373 (1981).
Winter, G., Fields, S., Brownlee, G. G.: Nucleotide sequence of the hemagglutinin gene of a human influenza virus H1 subtype. Nature *292,* 72–75 (1981).
Young, J. F., Palese, P.: Evolution of human influenza A viruses in nature: Recombination contributes to genetic variation of H1N1 viruses. P.N.A.S. *76,* 6547–6551 (1979).
Zhdanov, V. M., Zakstalskaya, L. Ya., Isachenko, V. I., Resnik, V. I., Andreyev, W. P., Lvov, D. K., Yakhno, M. A., Braude, N. A., Pysina, T. V., Podchernyaeva, R. Ya.: Return of epidemic A1 (H1N1) influenza virus. Lancet *I,* 294–295 (1978).

Subject Index

D. W. Kingsbury

Abbreviations used in this index

IAV, influenza A virus
IBV, influenza B virus
ICV, influenza C virus
IV, influenza virus

Special usages

above and *below* in cross-references indicate that the reference is to another subheading under the same heading, not to a separate heading.