Biochem. J. (1983) **211**, 281–294
Printed in Great Britain

REVIEW ARTICLE

Structure and function of the influenza virus genome

John W. McCAULEY and Brian W. J. MAHY
Division of Virology, Department of Pathology, University of Cambridge, Laboratories Block, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, U.K.

Introduction

The virus responsible for human epidemic influenza was first isolated 50 years ago by laboratory infection of ferrets with human nasal washings (Smith *et al.*, 1933). This isolation was the culmination of 15 years of research to find the causative agent of the influenza pandemic of 1918, which in the space of 4 months resulted in 20 million deaths, and since when epidemic influenza has remained the most serious unconquered acute threat to human health (Grist, 1979). Since 1918, influenza pandemics have occurred in 1957 (Asian influenza) and in 1968 (Hong Kong influenza), but between these episodes, epidemics of serious proportions occur almost annually. These epidemics and pandemics arise as a result of antigenic variation of two virion surface antigens, the haemagglutinin and neuraminidase, changes in the former being the more frequent and significant. The variation occurs either as a result of mutation and selection under immunological pressure, when it is known as 'antigenic drift', or by replacement of the entire gene encoding the haemagglutinin or neuraminidase with that of another virus strain, when it is known as 'antigenic shift'. The former process usually gives rise to new epidemics; the latter may result in a pandemic. In addition to man, influenza virus infects birds, horses and swine (Pereira, 1982) and has recently caused an outbreak of disease in harbour seals (Webster *et al.*, 1981). It is by genetic reassortment between human and other strains of influenza, particularly those infecting birds, that antigenic shift is thought to occur (Webster *et al.*, 1982). This reassortment is possible because of the peculiar structure of the influenza virus genome, in eight distinct single-stranded RNA pieces. Before considering the genome structure in detail, a brief description of the influenza virion will be presented.

The virion

The influenza genome RNA segments are associated in the mature virion with a basic, arginine-rich protein, comprised of 498 amino acids (56000 M_r) known as the nucleoprotein. There are about 1000 nucleoprotein molecules per virion (Inglis *et al.*, 1976). A number of electron microscopic studies have failed to reveal the exact structure of the ribonucleoprotein, although it seems likely that each genome RNA segment is separately encapsidated as a ribonucleoprotein, that the ribonucleoproteins have a helical left-handed configuration (Compans *et al.*, 1972; Murti *et al.*, 1980; Heggeness *et al.*, 1982), and that the encapsidated RNA is accessible to ribonuclease (Duesberg, 1969, 1970). Three types of influenza virus, A, B and C, can be defined on the basis of the antigenicity of the ribonucleoprotein. Complement-fixing antibodies raised against the ribonucleoproteins of all type A viruses cross-react, whereas no cross-reaction is observed with types B or C. The majority of influenza virus epidemics, and all pandemics so far recorded, are caused by influenza virus type A, which is also the type associated with influenza of avian, equine and swine species. Type B has only been isolated from human infections, and the same was true of the rarer type C virus until a recent Chinese investigation isolated 15 new strains of influenza virus C from swine (Guo *et al.*, 1983). Antigenic shift has so far only been recorded with type A influenza virus, and this review will concentrate on studies of type A viruses.

The ribonucleoprotein is surrounded in the virion by a shell of matrix protein which is hydrophobic and comprised of 252 amino-acids (28000 M_r) (Allen *et al.*, 1980; Winter & Fields, 1980; Lamb & Lai, 1981; McCauley *et al.*, 1982). There are about 2500 molecules of matrix protein per virion (Inglis *et al.*, 1976). Within the matrix shell, and associated in an undetermined manner with the genome RNA, are three high molecular weight P proteins, which form a RNA-dependent RNA polymerase complex capable of transcribing complementary mRNA molecules from the genome RNA template (Kawakami & Ishihama, 1983). Two of these P proteins, PB_1 and PB_2,. are basic and are comprised of 757 and 759 amino-acids respectively; the third, PA, is acidic, and comprised of 716 amino-acids (Horis-

Abbreviation used: vRNA, virus RNA.

berger, 1980; Winter & Fields, 1982; Fields & Winter, 1982). There are only about 15 molecules of each P protein in the virion (Inglis *et al.*, 1976).

Surrounding the matrix protein shell is a host-cell-derived lipid bilayer (Landsberger *et al.*, 1971) into which are inserted spike-like projections (Horne *et al.*, 1960; Hoyle *et al.*, 1961). The proteins comprising the surface projections were identified as glycoproteins after removal from the virion surface using Pronase, the plant proteinase bromelain, or detergent treatment (Compans *et al.*, 1970; Schulze, 1970, 1972; Brand & Skehel, 1972). Approx. 80% of the surface projections are haemagglutinin and the remaining 20% neuraminidase (Laver & Valentine, 1969; Webster *et al.*, 1982). The haemagglutinin is a trimer with an M_r of about 225000, consisting of three identical subunits each containing two polypeptide chains, HA_1 and HA_2, linked in each subunit by a single disulphide bond (Waterfield *et al.*, 1981; Ward & Dopheide, 1980). The neuraminidase is a tetramer, which appears in electron micrographs to be mushroom-shaped, of approx. M_r 200000 (Griffith, 1975; Wrigley, 1979; Blok *et al.*, 1982).

Both haemagglutinin and neuraminidase have been crystallized after bromelain or Pronase treatment respectively (Brand & Skehel, 1972; Laver, 1978; Wright & Laver, 1978). The three-dimensional structure of the haemagglutin trimer has recently been determined by X-ray diffraction (Wilson *et al.*, 1981).

The haemagglutinin and neuraminidase undergo post-translational modification by glycosylation (Klenk & Rott, 1980), but whereas the neuraminidase remains intact (Blok *et al.*, 1982), the haemagglutinin is processed by removal of an *N*-terminal signal sequence (McCauley *et al.*, 1979) and subsequently cleaved into HA_1 and HA_2 (Waterfield *et al.*, 1979).

Variation in the antigenicity of the haemagglutinin and neuraminidase is the basis of division of influenza type A viruses into subtypes. Human viruses isolated in the 1933–1956 era are designated H1N1; those of the 'Asian' type isolated from 1957–1967 are designated H2N2; those isolated since 1968 of the 'Hong Kong' type are designated H3N2. A new subtype is only introduced for major changes in antigenicity, and within any subtype there are numerous variants. In all, considering influenza A viruses of horses, swine and birds in addition to man, there are currently 13 H subtypes and 9 N subtypes (WHO Memorandum, 1980, Hinshaw *et al.*, 1982).

The RNA genome

It has become apparent that the influenza virus genome consists of eight single-stranded negative-sense RNA molecules which are present in the virus particle in near equimolar amounts. The eight RNA genome segments can be separated by polyacrylamide-gel electrophoresis (McGeoch *et al.*, 1976; Palese, 1977; Scholtissek, 1978) and range in size from 890 nucleotides (segment 8) to 2341 nucleotides (segments 1 and 2) (see below). Protein and mRNA coding functions have been mapped to individual segments in part by hybrid-arrested translation of mRNA *in vitro* (Inglis *et al.*, 1977, 1979; Lamb & Choppin, 1980, 1981) and by analysis of the RNA composition of recombinant viruses whose proteins have strain-specific differences such as electrophoretic mobility or antigenicity (reviewed by Barry & Mahy, 1979; Scholtissek, 1979; Palese, 1977). Each of the six largest RNA segments encodes a single polypeptide present in both the infected cell and the virion. The two smallest segments, 7 and 8, each encode two polypeptides, matrix protein and M_2, and NS_1 and NS_2 respectively; only matrix protein is found in the virion, the others are non-structural virus polypeptides (Inglis *et al.*, 1976, 1979; Lamb & Choppin, 1979, 1981). Each protein is translated from a single mRNA species. An additional (eleventh) species of polyadenylated capped RNA encoded by RNA segment 7 has been isolated from infected cells for which the peptide product has yet to be identified in the virus or infected cell (Inglis & Brown, 1981; Lamb *et al.*, 1981). All mRNA molecules are of opposite polarity to the genome RNAs and there is no evidence that any polypeptides are translated from the virion RNA sense.

Nucleotide sequences

The complete nucleotide sequence is now available for the RNA of one human strain of influenza virus, A/PR/8/34, which consists of 13588 nucleotides; the sequences of 45 individual segments of A and B strains of influenza have now been published representing 41 different RNA segments (Table 1). The most extensively studied gene and gene product is the haemagglutinin which is encoded by RNA segment 4, since it is against haemagglutinin that neutralizing antibody is directed and changes in the primary structure of the haemagglutinin are associated with epidemics of influenza. However, variation in these sequences will not be considered in detail here, as recent reports have considered the relevance of the changes (Wiley *et al.*, 1981; Skehel *et al.*, 1982; Webster *et al.*, 1982). The sizes of the segments and of the encoded polypeptides are summarized in Table 1.

The 5′ and 3′ terminal sequences of influenza A virus RNA segments are conserved between the segments and amongst different subtypes (Skehel & Hay, 1978; Robertson, 1979). In influenza virus types B and C, the 3′ and 5′-terminal nucleotide

Table 1. *Nucleotide sequences of influenza RNA*

Segment	Encoded polypeptide(s)	Strain	Length (nucleotides) of Segment	Length (nucleotides) of Coding sequence	Reference
1	PB_2	A/PR/8/34	2341	2277	Fields & Winter (1982)
		A/WSN/33	2341	2277	Kaptein & Nayak (1982)
		A/FPV/Rostock	2341	2277	I. J. Roditi & J. S. Robertson (unpublished work)
2	PB_1	A/PR/8/34	2341	2271	Winter & Fields (1982)
		A/WSN/33	2341	2271	Sivasubramanian & Nayak (1982)
		A/NT/60/68	2341	2271	Bishop *et al.* (1982*a*)
3	PA	A/PR/8/34	2233	2148	Fields & Winter (1982)
		A/NT/60/68	2233	2148	Bishop *et al.* (1982*b*)
		A/FPV/Rostock/34	2233	2148	J. S. Robertson *et al.* (unpublished work)
4	Haemagglutinin	Sub-type H1			
		A/PR/8/34	1778	1698	Winter *et al.* (1981*a*)
		A/WSN/33	1775	1695	Hiti *et al.* (1981)
		Sub-type H2			
		A/JAP/305/57	1733	1686	Gething *et al.* (1980)
		Sub-type H3			
		A/Duck/Ukraine/63	1765	1698	Fang *et al.* (1981)
		A/Aichi/2/68	NR*	1698	Verhoeyen *et al.* (1980)
		A/Memphis/102/72	1768	1701	Sleigh *et al.* (1980)
		A/Victoria/3/75	1768	1701	Min Jou *et al.* (1980)
		A/Bangkok/1/79	NR*	NR*	Both & Sleigh (1981)
		A/NT/60/68	1765	1698	Both & Sleigh (1980)
		A/England/21/77	1765	1698	Hauptmann *et al.* (1983)
		Sub-type H7			
		A/FPV/Rostock/34	1742	1689	Porter *et al.* (1979)
		B-type			
		B/Lee/40	1882	1752	Krystal *et al.* (1982)
5	Nucleoprotein	A/PR/8/34	1565 †	1494 †	Winter & Fields (1981) Van Rompuy *et al.* (1981)
		A/NT/60/68	1565	1494	Huddleston & Brownlee (1982)
6	Neuraminidase	Sub-type N1			
		A/PR/8/34	1413	1362	Fields *et al.* (1981)
		A/WSN/33	1409	1359	Hiti & Nayak (1982)
		Sub-type N2			
		A/NT/60/68	1467	1407	Bentley & Brownlee (1982)
		A/Udorn/72	1466	1407	Markoff & Lai (1982)
		A/Victoria/3/75	1467	1407	Van Rompuy *et al.* (1982)
		B-type			
		B/Lee/40	1557	NA: 1398 NA_2: 300¶	Shaw *et al.* (1982)
7	Matrix protein, M_2‡, M_3‡	A/PR/8/34	1027	M: 756 M_2: 291‡	Winter & Fields (1980)
			1027	M: 756 M_2: 291‡	Allen *et al.* (1980)
		A/Udorn/72	1027	M: 756 M_2: 291‡	Lamb & Lai (1981); Lamb *et al.* (1981)
		A/FPV/Rostock	1027	M: 756 M_2: 291‡	McCauley *et al.* (1982)
		B-type			
		B/Lee/40	1191	M: 744 M_2: 585§	Briedis *et al.* (1982)
8	Nonstructural (NS_1) Nonstructural (NS_2)‖	A/PR/8/34	890	NS_1: 690 NS_2: 363	Baez *et al.* (1980)
		A/PR/8/34	890	NS_1: 690 NS_2: 363	Winter *et al.* (1981*b*)
		A/Udorn/72	890	NS_1: 711 NS_2: 363	Lamb & Lai (1980)
		A/Duck/Alberta/60/76	890	NS_1: 690 NS_2: 363	Baez *et al.* (1981)
		A/FPV/Rostock/34	890	NS_1: 690 NS_2: 363	Porter *et al.* (1980)
		A/FM/1/47	890	NS_1: 606 NS_2: 363	Krystal *et al.* (1983)
		A/FW/1/50	890	NS_1: 711 NS_2: 363	Krystal *et al.* (1983)
		A/USSR/90/77	890	NS_1: 711 NS_2: 363	Krystal *et al.* (1983)
		B-type			
		B/Lee/40	1096	NS_1: 843 NS_2: 366	Briedis & Lamb (1982)

* Non-coding or signal sequence not reported.
† An error has been reported (see Huddleston & Brownlee, 1982).
‡ Derived from spliced RNAs and the size deduced from sequences of spliced mRNA of A/Udorn/72 (Lamb *et al.* 1981). M_3 peptide undetected.
§ Size of second largest open reading frame (peptide undetected).
¶ Size of second largest open reading frame: the NA_2 peptide has been detected (R. A. Lamb, personal communication).
‖ Derived from spliced RNA; the size is deduced from the sizes of A/Udorn/72 mRNAs (Lamb *et al.*, 1980).

sequences also appear to be conserved (Skehel & Hay, 1978; Desselberger *et al.*, 1980). For influenza A viruses, at the 5′ end of each segment the 13-nucleotide sequence 5′-AGUAGAAACAAGG is found, followed by a triplet unique to each RNA segment, and then a tract of oligo(U). At the 3′ end of each RNA segment a 12-nucleotide sequence of $_{3'HO}$UCGUUUUCGUCC is found, which is conserved in most RNA segments but found as $_{3'HO}$UCGCUUUCGUCC in segments 1–3, and in segment 7 of human influenza strains (Robertson, 1979; Allen *et al.*, 1980; Winter &

Fields, 1980). Apart from these sequences, no homologies have been reported between segments even for the two P genes (encoding PB_1 and PB_2) of identical nucleotide length of 2341 nucleotides. Whether the absence of homology means that the RNA segments were independently derived, or have undergone very extensive divergence during evolution from a common ancestor, is unknown.

Comparison of the nucleotide sequences of individual segments between different strains reveals a varying degree of homology. The most variable RNA segments are those coding for the surface glycoproteins. The haemagglutinin can show as little as 35% amino acid sequence homology between sub-types with insertions and deletions included to maximize any homology; such low homology can be detected in a comparison of A/PR/8/34 from H1 and for example A/Aichi/2/68 from H3 (Winter *et al.*, 1981*a*). Between other haemagglutinin sub-types around 40–45% amino acid sequence homology is found: H2 versus H3 (40%), H2 versus H7 (42%) (Gething *et al.*, 1980). However 69% homology is found in comparison of H1 and H2 which presumably reflects divergence from a more recent common ancestor (Winter *et al.*, 1981*a*). Within a subtype, for example H3, there is much more homology: 92% homology in amino acid sequence is seen between viruses isolated in 1968 (A/NT/60/68) and 1979 (A/Bangkok/1/79), and most of the changes occur in the HA_1 polypeptide chain (Both & Sleigh, 1981) the region which contains the antigenic sites of the molecule (Skehel *et al.*, 1982).

The other external glycoprotein, the neuraminidase, shows a similar degree of variation between subtypes; no more than 41% amino acid sequence homology is found between N1 and N2 (Bentley & Brownlee, 1982; Markoff & Lai, 1982; Van Rompuy *et al.*, 1982).

Variability in the nucleotide and polypeptide sequences of genes which encode internal proteins has also been estimated. For nucleoprotein, the antigenicity of which provides the basis for distinguishing the A and B viruses, only three reports of RNA sequences have been published, and two of these are from A/PR/8/34. The sequences are the same in length (one of the PR8 sequences contained an error, quoted in Huddleston & Brownlee, 1982), and show 94% amino acid homology and 92% nucleotide sequence homology. The significance of the close homology between these viruses is unclear since it is not known whether nucleoproteins from these viruses can be distinguished antigenically (van Wyke *et al.*, 1980; Schild *et al.*, 1979).

On the basis of RNA hybridization data, Scholtissek & von Hoyningen-Huene (1980) proposed three groups within the A viruses which are defined by the genetic relatedness of RNA segment 8. Seven reports of the nucleotide sequence of segment 8 from A type viruses have, to some extent, borne this out. For example, the nucleotide sequence of A/Duck/Alberta/60/76 shows only 73% homology between it and A/PR/8/34 (Baez *et al.*, 1981) which is in agreement with the proposed groups. However, the sequences of A/PR/8/34 (Baez *et al.*, 1980; Winter *et al.*, 1981*b*), A/Udorn/72 (Lamb & Lai, 1980) and A/FPV/Rostock/34 (Porter *et al.*, 1980) were all closely related, showing about 90% nucleotide sequence homology, a result which was unexpected from the hybridization data.

A higher degree of nucleotide sequence homology has been found for RNA segment 7 from both human and avian sources (Allen *et al.*, 1980; Winter & Fields, 1980; Lamb & Lai, 1981; McCauley *et al.*, 1982). A 91–92% nucleotide sequence homology of segment 7 is reflected in a 97% amino acid sequence homology in matrix protein and approximately 93% amino acid sequence homology between the M2 gene products. These results suggest that matrix protein is highly conserved due to an undefined selective pressure.

Three nucleotide sequences are now available for each of the largest RNA segments (Table 1). Within each segment there is approx. 85–90% homology in nucleotide sequence and 90–95% homology in amino acid sequence, with the exception of WSN and PR8 segment 1: these show 95% homology in both nucleotide and amino-acid sequences.

In summary, it seems that different strains show remarkably high degrees of nucleotide sequence homology. Deletions and/or insertions of sequence are rare and occur only in triplets. The largest insertion found to date occurs in the neuraminidase gene, where an insertion of 45 nucleotides can be seen between examples of the N1 and N2 sub-type neuraminidases. An insertion of 12 nucleotides at the HA_1–HA_2 connecting peptide can be seen in the haemagglutinin of FPV/Rostock. Most insertions or deletions are found in the glycoprotein genes and encode a variable number of amino acids.

Some attempts have been made to estimate the rate of evolution of influenza virus genes, but such estimates may be fallible. Estimates of the speed of evolution rely on the assumption that the date of isolation of a virus represents the time at which divergence started: it is however quite feasible that divergence started at a later date, when the virus came under selective pressure. Furthermore, since influenza viruses show a high degree of genetic reassortment one gene may not have evolved from the previous isolate but from a different virus, and in such circumstances the determination of mutation rate may be misleading. Nevertheless, analysis of the amino acid sequence and its variation may be of interest. Huddleston & Brownlee (1982) have estimated a rate of 1% amino acid sequence change per year for the haemagglutinin between 1968

(A/NT/60/68) and 1979 (A/Bangkok/79), whereas the rate of change of other genes was much slower, varying from 0.07% amino acid changes per year for matrix protein, to 0.18% amino acid changes per year for nucleoprotein and up to 0.31% per annum for NS_1. These ratios were estimated by a comparison of viruses isolated in 1934 with those in 1968 or 1972, but more examples are needed to provide a reliable estimate of the rate of evolution of these genes. The reported sequences of the neuraminidase subtype N2, however, provide another opportunity to measure the rate of evolution from 1968 to 1972 and 1975 (Bentley & Brownlee, 1982; Markoff & Lai, 1982; van Rompuy *et al.*, 1982). The figures for the rate of evolution of neuraminidase are all close to 1% amino acid changes per annum (close to the rate of evolution of the haemagglutinin during the same period). The observation that most amino acid sequence changes once established remain in the molecule and do not revert, suggests that these three neuraminidase genes lie on a common evolutionary pathway. Further study of the N2 sub-type from 1957 to the present day may reveal whether 1% amino acid change per year has been tolerated for 26 years, the period during which the N2 subtype has existed in the human population. This will involve determination of the sequence of neuraminidase genes from N2 viruses isolated between 1957 and 1968 from H2N2 viruses. When the complete nucleotide sequences of RNA segments of many more virus isolates are known it may be possible to draw evolutionary trees of viruses and individual genes, as has been initiated for the glycoprotein genes based on 3′-terminal sequences (Air, 1981; Blok & Air, 1982).

Structure of intracellular virus-specific RNAs

In the infected cell the eight vRNA segments direct the synthesis of two forms of RNA of positive polarity, capped polyadenylated mRNA molecules and uncapped non-polyadenylated cRNA molecules, the presumed templates for vRNA synthesis. Characterisation of the terminal sequences of vRNA, cRNA and mRNA species has been a subject of intensive investigation in order to define the relationships between the virus genome, the polynucleotide products it encodes, and the host cell.

mRNA structure

The study of the structure of capped and polyadenylated transcripts of the influenza genome has intensified recently following the discovery that influenza mRNA transcription can be primed *in vitro* by the addition of capped RNA molecules to detergent-solubilized virus (Bouloy *et al.*, 1978). Influenza mRNAs synthesized *in vitro* using globin mRNA, reovirus mRNA or alfalfa mosaic virus RNA to prime synthesis contain, at their 5′ ends, capped non-influenza virus encoded sequences approx. 10–15 nucleotides in length derived from the 5′ end of the primer RNA molecules (Bouloy *et al.*, 1978, 1979; Robertson *et al.*, 1980; Krug, 1981). Determination of the sequences of mRNA molecules isolated from infected cells also revealed heterogeneous non-influenza sequences, 10–15 nucleotides in length, at their 5′ termini; this was demonstrated by analysis of cloned DNA copies of individual mRNAs (Dhar *et al.*, 1980; Caton & Robertson, 1980; Beaton & Krug, 1981) or by using restriction enzyme fragments isolated from cloned DNA copies of viral genes to prime reverse transcription of viral mRNA (Caton & Robertson, 1980; Lamb & Lai, 1980; Lamb *et al.*, 1981). These experiments also indicated that priming by capped host cell RNAs, or RNAs added during transcription *in vitro*, initiates polymerization at either the first or second 3′-terminal nucleotide of vRNA.

Both mRNA synthesized *in vitro* and infected cell mRNA are polyadenylated at their 3′ ends. These mRNAs have been shown to lack sequences complementary to the 5′-terminal 16 nucleotides of vRNA by nuclease protection experiments, which also defined the site of poly(A) addition *in vivo* as an oligo(U) tract immediately 3′ to the 5′ terminal 16 nucleotides of vRNA (Skehel & Hay, 1978; Robertson *et al.*, 1981). The polypeptide coding sequences are all totally encoded by influenza-derived RNA and the largest eight viral mRNA molecules isolated from the infected cell are almost certainly collinear transcripts of vRNA (Hay *et al.*, 1977*a*,*b*). However there are three smaller virus-specific RNA species found in the infected cell (Inglis & Brown, 1981) which are capped and polyadenylated but are not collinear transcripts of the influenza genome segments (Lamb & Lai, 1980; Lamb *et al.*, 1981). These RNA species, which are derived from RNA segments 7 and 8, are not found as products of transcription *in vitro* and are thought to be processed *in vivo* by splicing of mRNAs complementary to RNA segments 7 and 8 (Fig. 1). In support of this, the nucleotide sequences found at the 'splice-junctions' of these RNA molecules are similar to the 'consensus' sequences found at splice sites in other eukaryotic mRNA molecules (Mount, 1982), although definite evidence to indicate that solely host cell mechanisms are sufficient to produce these three RNA species remains to be published. The nucleotide sequences of these RNA molecules from cells infected with the A/Udorn/72 strain show that segment 7 encodes two capped polyadenylated 'spliced' RNA species, 318 and 289 influenza-virus-specific nucleotides in length (Lamb *et al.*, 1981). The RNAs are produced by the combination of a single splice acceptor site at nucleotide 740 of

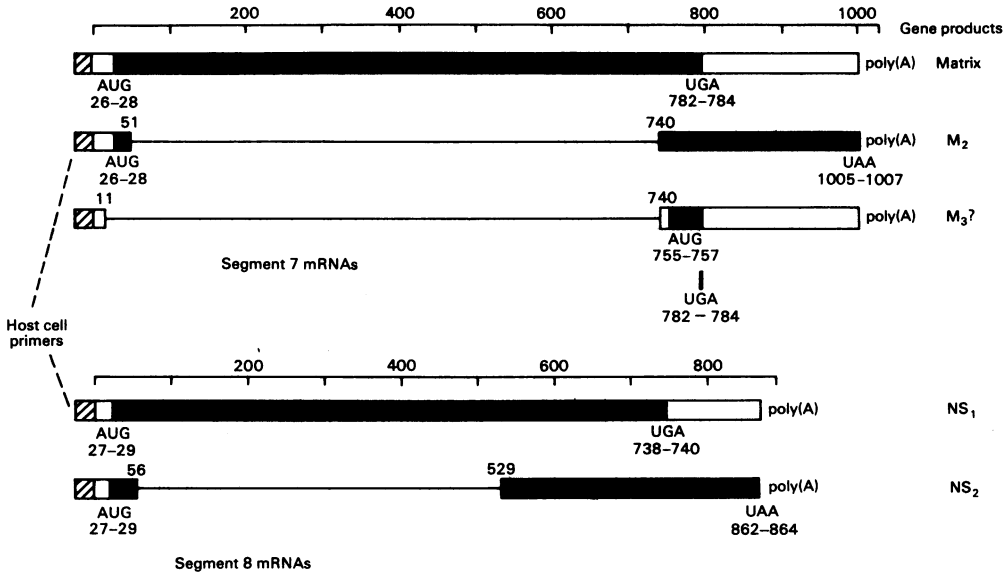

Fig. 1. *Spliced mRNA products complementary to segments 7 and 8 from A/Udorn/72 (Lamb & Lai, 1980; Lamb et al., 1981)*

The heterogeneous host cell primers are indicated by cross hatching, the coding sequences by filled boxes, and the non-coding sequences by open boxes. The regions of RNA removed by splicing are indicated by thin lines. The numbers are the nucleotide positions from the 3′ end of influenza virion RNA. The initiation and termination codons are indicated.

cRNA and two splice donor sites at nucleotide positions 11 and 51. The three predicted polypeptides encoded by RNA segment 7 are matrix protein, which is encoded by a single reading frame with an initiation codon at nucleotide 26–28 and termination at 782–784; M_2 protein with the same initiation codon but a splice to join nucleotide 51 to 740 resulting in removal of 689 nucleotides, and a resultant change of reading frame terminated at nucleotides 1005–1007; and M_3 peptide, the mRNA for which has lost the initial AUG by a splice which results in removal of nucleotides 12–739 inclusive: the first AUG codon at position 755–757, acting as initiator, is in the same reading frame as matrix protein so that termination occurs at 782–784. M_3 peptide, therefore, would represent the *C*-terminal nine aminoacids of matrix protein (see Fig. 1). The small poly(A)-containing RNA derived from segment 8 encodes NS_2 polypeptide (Stephenson *et al.*, 1977; Inglis *et al.*, 1979; Lamb & Choppin, 1979, 1980) and results from a single splice between nucleotide 56 and 529 so that NS_1 and NS_2 are both initiated at nucleotides 27–29 and have a common sequence of nine amino acids at the *N*-terminus. NS_1 and NS_2 are read in different reading frames following the splice for 70 amino acid residues; NS_1 then terminates at nucleotides 738–740, NS_2 at 862–864 (Lamb & Lai, 1980). Although NS_2 and M_2 have been shown both *in vivo* and by translation *in vitro*, no protein sequence analysis has confirmed the predicted initiation and termination sites and M_3, the *C*-terminal peptide of matrix protein, has yet to be found by translation *in vitro* or *in vivo* in the virion, or the infected cell.

The signals for polyadenylation of mRNA *in vitro* and *in vivo* are not fully understood. Only capped RNAs are found to be polyadenylated in the cell and the mechanism may be similar to that envisaged in vesicular stomatitis virus (VSV) transcription, where it is thought that the RNA polymerase may 'stutter' at the oligo(U) tracts, thus transcribing the poly-(A) tract (Schubert *et al.*, 1980; Robertson *et al.*, 1981); poly(A) signals such as have been identified for eukaryotic mRNAs transcribed by RNA polymerase II (Proudfoot, 1982) are not present in the virion RNAs. Whether the sequence complementary to the 5′-terminal 16 nucleotides of vRNA is present 3′ to the poly(A) tract at early stages of transcription (and subsequently removed) remains to be established.

Although there are other possible open reading frames encoded in influenza mRNA or even present in the negative strand, no evidence has yet been produced to suggest that other mRNA molecules exist or that any more gene products are synthesized *in vivo*.

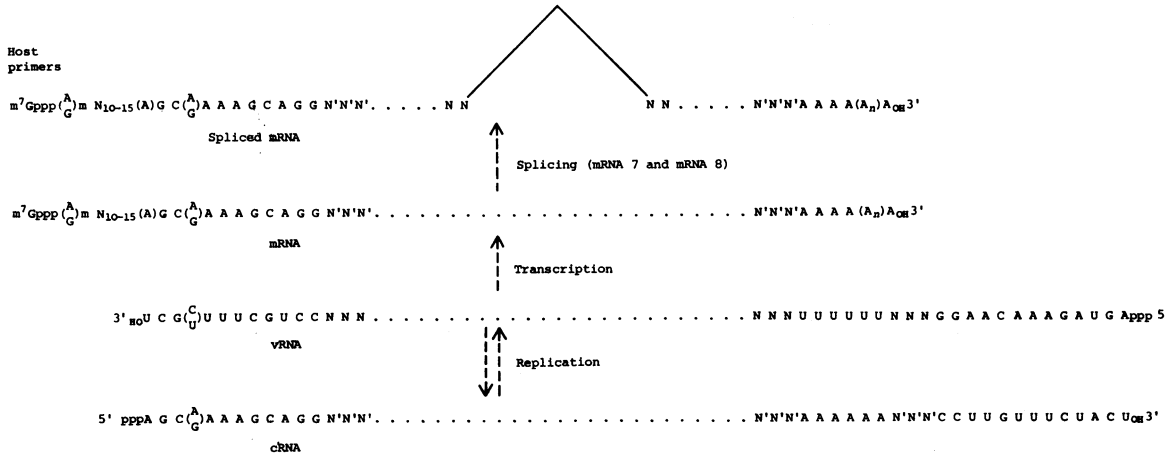

Fig. 2. *Relationships between the classes of RNA found in the infected cell (modified from Hay et al.*, 1980) Details of the splicing events are presented in Fig. 1.

Nonpolyadenylated complementary RNA

In addition to mRNAs, non-polyadenylated RNA transcripts complementary to the eight virus genome RNAs can be found in the infected cell. On the basis of nuclease protection experiments and polyacrylamide-gel mobility these cRNAs were shown to be complete transcripts of the genome RNA (Hay *et al.*, 1977*a,b*, 1978). More recently the 3′- and 5′-terminal sequences of these RNAs (Hay *et al.*, 1982) have confirmed the nuclease protection data, and show that their mode of synthesis differs from mRNA in that priming using host cell RNAs does not occur; the 5′-terminal nucleotide is pppAp and the 3′-terminal sequence is complementary to the 5′ terminus of vRNA. Full length cRNA molecules have not been isolated from the products of transcription *in vitro* (Hay & Skehel, 1979; Rochovansky, 1981); indeed in the cell, synthesis of cRNA molecules requires continued protein synthesis (Barrett *et al.*, 1979; Mark *et al.*, 1979; Hay *et al.*, 1980). It is evident therefore that the 5′- and 3′-terminal sequences of vRNA are conserved and complementary to those of cRNA. The role of these terminal sequences of both cRNA and vRNA has yet to be determined. Any similarity in the sequences at the 5′ ends of vRNA and cRNA may reflect similar recognition sites for polypeptides in the assembly of functional vRNA or cRNA ribonucleoprotein complexes effecting control of transcription and replication.

In the cell, the cRNA molecules are not found associated with polysomes and therefore are thought not to play a role in protein synthesis although they will direct polypeptide synthesis *in vitro* (Stephenson *et al.*, 1977; Inglis & Mahy, 1979).

The characteristics of cRNA clearly suggest that its role *in vivo* is as template for new vRNA synthesis.

vRNA

The characteristics of vRNA in the host cell are generally assumed to be the same as virion RNA. Young & Content (1971) determined the degree of phosphorylation at the 5′ terminus as pppAp (the same as cRNA; Hay *et al.*, 1982). This implies independent initiation of synthesis of individual strands and not processing from a single transcript. A summary showing the inter-relationships between the influenza RNAs found in the infected cell is given in Fig. 2.

Independent initiation of transcription and replication

Although independent initiation of replication and transcription is implied by the terminal structure of the virus-specific RNAs found in the infected cell the possibility exists of an ordered sequence of transcription and replication of the virion RNA segments. The sensitivity of RNA synthesis to u.v. light has been utilized to estimate the relative target sizes for transcription (Abraham, 1979; Pons & Rochovansky, 1979) and replication (Smith & Hay, 1982). In all cases the size of the target was proportional to the size of the RNA species analysed, a result consistent with independent initiation and a lack of ordered sequence of synthesis of segments. No evidence has been presented on the u.v.-target size for the non-collinear poly(A)-containing RNAs derived from segments 7 and 8.

The sites of synthesis of virus-specific RNAs

It has long been established that influenza virus requires a functional host cell nucleus for replication (Barry *et al.*, 1962; Mahy, 1970; Follett *et al.*, 1974; Mahy *et al.*, 1980) and more specifically DNA-dependent RNA polymerase II activity (Mahy *et al.*, 1972; Spooner & Barry, 1977; Lamb & Choppin, 1977). Analysis of influenza RNA by using radioactive probes showed that early in infection, a relatively high proportion of virus-specific RNA could be detected in the nucleus of infected cells, and these RNAs were found to accumulate in the nucleus in the presence of certain drugs (Barrett *et al.*, 1979; Mahy *et al.*, 1981; Mark *et al.*, 1979). The discovery of host cell RNA sequences at the 5′ termini of influenza virus mRNA (Dahr *et al.*, 1980; Caton & Robertson, 1980) renewed efforts to define the site of mRNA synthesis. By rapid pulse labelling (2 min pulse) of BHK21 or CEF cells with [^{3}H]uridine and the use of a non-aqueous method for nuclear purification, Herz *et al.* (1981) demonstrated that a large proportion (80–90%) of the virus-specific pulse-labelled mRNA could be found in the nuclear fraction which was lost from CEF cells using normal aqueous procedures. The results of previous attempts to locate the site of synthesis of rapidly labelled RNA, which concluded that transcription was cytoplasmic, were presumably erroneous for this reason (Hay *et al.*, 1977*b*). Using an aqueous or non-aqueous procedure with BHK-21 cells evidence was also obtained that poly(A)-free cRNA synthesis was nuclear (Herz *et al.*, 1981). By a different procedure Jackson *et al.* (1982) found pulse-labelled influenza mRNA associated with the nuclear cage, but their results suggested that vRNA was also nuclear. The results with vRNA synthesis require further study or confirmation using non-aqueous nuclear separation and hybridization to suitably purified influenza mRNA. Nuclear sites for transcription and replication are consistent with the finding that several virus-specific proteins (PB_1, PB_2, NP and NS_1) migrate into the nucleus after synthesis in the cytoplasm (Taylor *et al.*, 1970; Krug & Etkind, 1973; Krug & Soeiro, 1975; Briedis *et al.*, 1981).

Ribonucleoprotein structures and their properties

In the virion the eight single-stranded genome RNA molecules are associated with virus polypeptides in ribonucleoprotein complexes. These complexes can be separated from other virus-specific polypeptides by centrifugation on density gradients formed from sucrose, glycerol, CsCl or Renografin (a methyl-glucamine salt of 3,5-diacetamido-2,4,6-tri-iodobenzoic acid) (Pons, 1971; Krug, 1972; Inglis *et al.*, 1976; Rochovansky, 1976; Krug & Soeiro, 1975; Caliguiri & Gerstein, 1978) or by gel electrophoresis (Duesberg, 1969; Rees & Dimmock, 1981). The ribonucleoprotein complexes contain PA, PB1 and PB2, nucleoprotein and a variable amount of matrix protein dependent on the method employed for their isolation. In infected cells, influenza virus RNAs of both positive and negative sense have been found as ribonucleoprotein complexes (Pons, 1971, 1976; Hay *et al.*, 1977*b*; Krug & Soeiro, 1975). Analysis of the RNA components of ribonucleoprotein complexes has shown that the polysome fraction contains only mRNA molecules and that most mRNA molecules are found in the polysome fraction (Glass *et al.*, 1975; Hay *et al.*, 1977*b*), whereas cRNA is found in the ribonucleoprotein fraction with sedimentation coefficients between 30 and 80S dependent on the size of the associated RNA molecules, very little being detected as free RNA (Hay *et al.*, 1977*b*).

Properties of virion ribonucleoproteins

The enzymatic properties of virion ribonucleoproteins, usually studied as detergent-disrupted virus, have been the subject of much interest due to their ability to synthesize mRNA *in vitro* using dinucleotides (McGeoch & Kitron, 1975; Carroll *et al.*, 1975; Rochovansky, 1976; Hay & Skehel, 1979; Plotch & Krug, 1977, 1978), or capped RNAs (Bouloy *et al.*, 1978, 1979), as primers of transcription. Through the use of crosslinking by u.v. light or photoreactive cap analogues, PB_2 in the transcribing virion ribonucleoprotein complex was shown to bind to RNA molecules containing a methylated cap 1 structure [$m^7G(5')ppp(5')Nm$] and so to be responsible for one of the steps in priming influenza mRNA synthesis by capped mRNAs (Ulmanen *et al.*, 1981; Blaas *et al.*, 1982*a*,*b*; Penn *et al.*, 1982). PB_1 has been shown to be involved in initiation of transcription of mRNAs: in the presence of viral ribonucleoproteins, capped mRNA and α-^{32}P-labelled GTP as the only nucleotide, PB_1 formed a covalent complex with the ^{32}P label under u.v. irradiation; this complex was not formed if the mRNA had its cap removed by β-elimination (Ulmanen *et al.*, 1981). Further evidence that PB_1 is involved in initiation of transcription came from experiments using dinucleotide primers and [α-^{32}P]CTP as the only mononucleotide; under such conditions a covalent complex between PB_1 and [α-^{32}P]CTP could be detected without the necessity for u.v. irradiation (Horisberger, 1982).

Both these results lend support to studies with temperature-sensitive (*ts*) mutants which revealed a role for PB_2 in stimulation of transcription *in vitro* by globin mRNA (Mahy *et al.*, 1981; Nichol *et al.*, 1981) or in cleavage of alfalfa mosaic virus RNA by a cap-dependent endonuclease (Ulmanen *et al.*,

1983). A role for PB_1 in transcription has also been shown using *ts* mutants (Mowshowitz, 1981; Horisberger, 1982). The results either from biochemical evidence or from *ts* mutants cannot rule out the possibility that virus specific proteins other than PB_2 or PB_1 co-operate in the transcriptase complex in effecting efficient cap-dependent priming, endonucleolytic activity or initiation of transcription.

Further purification of individual polypeptide components of the genome ribonucleoprotein complex is clearly warranted, and some progress has been made to this end by Kawakami & Ishihama (1983) who managed to purify a ribonucleoprotein consisting of PA, PB_1 and PB_2 associated with vRNA which was free from nucleoprotein. This complex can transcribe vRNA in the presence of 3′-terminal complementary dinucleotide primers, but characterization of the initiation and termination sites and length of the RNA produced has yet to be published, and it is not clear whether the complex is stimulated by capped RNA molecules.

The mRNAs produced by solubilized virus or purified cores are incomplete transcripts of the genome (Hay *et al.*, 1978; Hay & Skehel, 1979; Rochovansky, 1981) and no full-length 'cRNA-like' molecules are produced *in vitro* (Hay & Skehel, 1979). The composition of the *in vitro* transcripts shows that all RNA segments are transcribed (McGeoch & Kitron, 1975; Plotch & Krug, 1977; Hay & Skehel, 1979; Skehel & Hay, 1978; Rochovansky, 1981) although the relative proportions of the transcripts produced by detergent-disrupted virus could be varied somewhat by a variation in the concentration of dinucleotide primers (Hay & Skehel, 1979). Whether a similar phenomenon occurs with purified ribonucleoproteins or when capped mRNA molecules are used as primers remains unreported.

Infected-cell ribonucleoproteins

Ribonucleoprotein structures which contain RNA polymerase activity are found in association with the microsomal as well as the nuclear cell fraction (Scholtissek & Rott, 1969; Skehel & Burke, 1969; Mahy *et al.*, 1975); the products of these complexes consist of positive-stranded RNA (Scholtissek, 1969; Mahy & Bromley, 1970; Hastie & Mahy, 1973) but their role in replication is unclear. All three P proteins are found in association with nucleoprotein in such enzymically active structures (Inglis *et al.*, 1976). Ribonucleoprotein complexes from both the cytoplasm and nucleus can be isolated by using CsCl gradients provided they are fixed with glutaraldehyde (Krug, 1971, 1972; Krug & Soeiro, 1975). Caliguiri & Gerstein (1978) identified two subclasses in influenza virus ribonucleoproteins from infected cells which differed in density in Renografin gradients. One subclass had identical composition to virion ribonucleoproteins; the other class, which was thought to represent replication ribonucleoproteins, was enriched in P proteins and found early in infection. Neither of these classes of ribonucleoprotein were active in transcription assays after purification, so their role in replication remains unclear, especially since the RNA sense of the ribonucleoproteins was not determined nor was it established whether one class represented an intermediate in ribonucleoprotein assembly.

The synthesis of individual ribonucleoproteins in the cell has been analysed after separation of five size classes of ribonucleoprotein with polyacrylamide gels (Rees & Dimmock, 1981). The assembly of individual ribonucleoproteins appears to be separately controlled since differences in total accumulation and nuclear : cytoplasmic ratios between ribonucleoprotein size classes could be detected (Rees & Dimmock, 1982).

At present replication complexes able to synthesize vRNA or cRNA *in vitro* have not been isolated, and progress in this area may require the establishment of conditions for coupled transcription and translation of suitable extracts from the infected cell, as recently accomplished for vesicular stomatitis virus (Davis & Wertz, 1982; Hill *et al.*, 1981; Ghosh & Ghosh, 1982).

Selective transcription and replication *in vivo*

It has been evident for some time that influenza virus exhibits temporal control of polypeptide synthesis (Skehel, 1972, 1973; Meier-Ewert & Compans, 1974; Inglis *et al.*, 1976; Lamb & Choppin, 1976), and the mechanism of this control has been the subject of investigation by a number of techniques. Analysis of the accumulation of individual RNA segments in infected cells has been performed using labelled hybridization probes (Mark *et al.*, 1979; Barrett *et al.*, 1979) or by translation *in vitro* of mRNA isolated from infected cells (Inglis & Mahy, 1979). Analysis of influenza RNA synthesis has also been carried out by pulse-labelling infected cells with [^{3}H]uridine followed by competition of the RNA of one sense with an excess of unlabelled vRNA or mRNA, nuclease digestion of the hybrids, and separation of the resulting double-stranded RNAs by polyacrylamide-gel electrophoresis (Hay *et al.*, 1977*b*; Smith & Hay, 1982). The conclusions, measuring synthesis or accumulation, were that temporal control of transcription to form mRNA occurs in most virus–cell systems, and that the levels of protein produced are a direct reflection of the level of mRNA in the cell (Inglis & Mahy, 1979; Hay *et al.*, 1980).

All the results from fowl plague virus-infected

chick embryo fibroblast cells indicate that there are three phases in the replication of the influenza genome (Hay *et al.*, 1977*b*; Inglis & Mahy, 1979; Barrett *et al.*, 1979). The first phase, which is independent of protein synthesis (primary transcription), results in the production of mRNA complementary to each segment of RNA in close to equimolar amounts (Hay *et al.*, 1977*b*; Barrett *et al.*, 1979). The second phase (early secondary transcription) is dependent on protein synthesis; template cRNA synthesis reaches a maximum at 30–90 min post-infection and is accompanied by amplified synthesis of mRNA and vRNA of segments 5 and 8. In the third phase (late secondary transcription), the mRNAs and vRNAs of segments 4, 6 and 7 are amplified and mRNA, cRNA and vRNA for segment 8 are reduced. mRNA and vRNA synthesis of segment 5 remains high throughout secondary transcription. Whilst mRNA synthesis for segments 1–3 remains low or unamplified throughout secondary transcription, vRNA segments 1–3 are amplified, although to a lower extent than the other segments (Hay *et al.*, 1977*b*; Barrett *et al.*, 1979; Smith & Hay, 1982). Thus with the exception of segments 1–3 the relative levels of synthesis of vRNA and mRNA of individual segments are similar during infection and occur in approximately equimolar amounts, which may imply that newly synthesized vRNA is only transcribed for a limited period (Smith & Hay, 1982).

Differential inhibition of mRNA and vRNA can be observed by the addition of cycloheximide or actinomycin D during infection. Cycloheximide abolishes vRNA and cRNA synthesis but the pattern of mRNA synthesis remains invariant from the time of addition (Barrett *et al.*, 1979; Hay *et al.*, 1980; Smith & Hay, 1982). Actinomycin D (2–5 μg/ml) reduces mRNA synthesis by greater than 90% (Barrett *et al.*, 1979; Mark *et al.*, 1979), but after its addition vRNA synthesis and cRNA synthesis remain invariant (Staroff & Bukrinskaya, 1981; Smith & Hay, 1982).

The correspondence between vRNA and mRNA synthesis suggests that the two events may be closely coupled. The time at which cRNA is synthesized suggests that the template for synthesis of mRNA in secondary transcription is newly synthesized vRNA, and that cRNA is synthesized only from input genome vRNA. Temperature-sensitive mutants have provided some evidence to support these notions. For example, cRNA synthesis may remain normal at the restrictive temperature in the absence of vRNA synthesis (Wolstenholme *et al.*, 1980; Mowshowitz, 1981; Thierry & Danos, 1982; Mahy, 1983) and secondary mRNA transcription may be dependent on synthesis of vRNA (and hence cRNA) in certain mutants (Ghendon *et al.*, 1982; Mahy, 1983).

In other cells the characteristics of influenza virus infection can differ (Lamb & Choppin, 1978; Bosch *et al.*, 1978). For example, in L cells no amplification of segment 7 mRNA or vRNA occurs, yet an amplification of cRNA 4 is reflected in an elevated synthesis of vRNA and mRNA 4 (Bosch *et al.*, 1978; Smith & Hay, 1982). In HeLa cells no reduction in synthesis of vRNA or mRNA for segment 8 is seen late in secondary transcription (Smith & Hay, 1982). Both these infections are non-productive, and yet vRNA synthesis and mRNA synthesis remain closely coupled.

A scheme (based on schemes proposed by Hay *et al.*, 1980; Smith & Hay, 1982; Ghendon *et al.*, 1982) can be put forward to show the events in influenza genome replication and is shown in Fig. 3. This proposes that the input genome is used as template not only for all primary transcripts, but

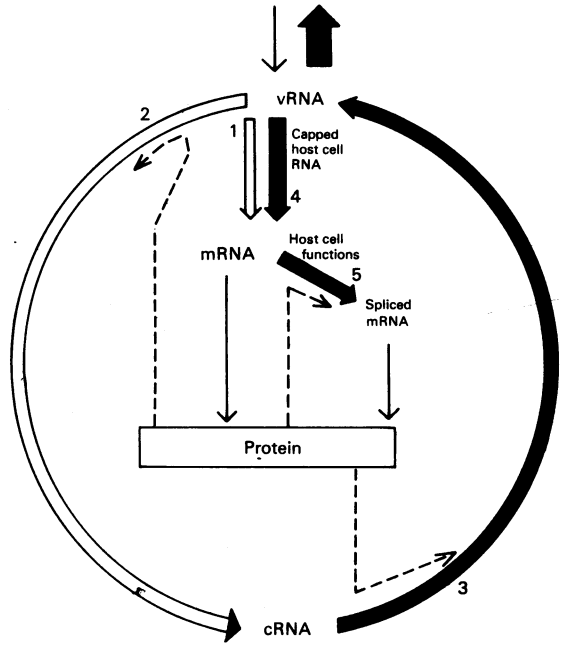

Fig. 3. *Scheme for influenza virus genome replication* ⇨, Transcription from input genome RNA; ➡, transcription from cRNA or newly synthesized vRNA; ----→ necessary interactions. Step 1: primary transcription, which occurs in the absence of protein synthesis but requires host cell transcription. Step 2: cRNA synthesis from input vRNA synthesis directed by cRNA, both require ongoing protein synthesis. Step 4: transcription of newly synthesized vRNA to form mRNA requires host cell caps. Step 5: modification of mRNAs 7 and 8 by splicing requires protein synthesis and undefined host cell functions.

also for all cRNA synthesis. mRNA is formed during secondary transcription only from vRNA synthesized *de novo*, and it is at the level of cRNA→vRNA that amplification and selective transcription occurs. Selective transcription of vRNAs→mRNAs only occurs for segments 1–3, and modification of transcripts occurs only with segments 7 and 8; furthermore, mRNA synthesis is not amplified by transcription of vRNA. The scheme does not define the mechanisms of selective transcription, but clearly control may be effected by the concentration of particular influenza virus protein(s) free in the infected cell or by the ratios of two influenza virus-specific polypeptides for example. Control of mRNA synthesis by the host cell has been explained, at least in part, by the need for capped RNA at the site of transcription, the nucleus, but how irradiation of cells with u..v. light prior to infection modifies the progression of influenza virus replication (Mahy *et al.*, 1977) remains unclear. Thus although the proposed scheme would support the view that the temporal control of transcription and polypeptide synthesis resides at the level of replication (Hay *et al.*, 1980; Smith & Hay, 1982; Ghendon *et al.*, 1982), further modification of transcription is required for the mRNAs of segments 1–3 and the spliced mRNAs of segments 7 and 8.

Assembly of the genome

The method of assembly of genome ribonucleoproteins into viral particles is not clearly understood. It has been proposed (Hirst, 1962) that assembly may be a random process but the current evidence is at variance with this proposal. Non-random segregation of parental genes in influenza virus recombination has been detected by Lubeck *et al.* (1979), suggesting that interaction between different RNA segments is involved in assembly and during some mixed infections more efficient interaction between segments of the homologous strain than with the heterologous strain may effect genetic linkage. Studies of incorporation of radiolabelled vRNA segments into progeny virus particles are hindered by the difficulty of pulse–chase labelling with uridine. Nonetheless, by labelling from 0 to 4 h post-infection and analysis of vRNA both in virus particles and in the infected host cells at 10 h, equimolar amounts of segments 4–8 were found in released virus, whereas in the cell, segments 5, 7 and 8 predominated and in addition, segments 1–3 were under-represented in the progeny virus (Smith & Hay, 1982). The reduction of segments 1–3, especially segment 3, may reflect the presence of incomplete RNAs in the infectious virus used in the experiment (Hay *et al.*, 1977; Skehel & Hay, 1978), and the high multiplicity of infection (Carter & Mahy, 1982). Parenthetically, late in infection a small cRNA related at its 3′ end to segment 3 has been found in cells infected with this virus (J. W. McCauley & A. J. Hay, unpublished results; Hay *et al.*, 1977*b*). These results show, therefore, that selective packaging of the genome of influenza virus occurs, which may be an important factor in recombination in nature. The mechanism by which selection occurs is not understood.

Conclusions

The replication of influenza virus involves a unique relationship with the host cell. Influenza virus mRNAs are synthesized in the cell nucleus, where they acquire capped 5′ termini from cellular RNA transcripts; two of the influenza mRNA species are further modified by splicing (presumably by host cell enzymes) which allows overlapping regions of the genome to direct synthesis of additional polypeptides. Temporal control of polypeptide synthesis can be seen in influenza virus-infected cells: this control is largely exerted at the level of replication, in particular synthesis of vRNA from its template, although control can also occur at the level of cRNA and mRNA synthesis and modification of mRNA by splicing. Temporal control of synthesis of individual gene products in this manner is possible due to the segmented nature of the genome.

The segmented genome also allows a high degree of genetic recombination (reassortment of genes) between influenza viruses. This reassortment plays an important role in the initiation of influenza epidemics throughout the world, since it can result in expression on the virion surface of new haemagglutinin and neuraminidase antigens (antigenic shift) against which the population has no immunity.

The emergence of nucleotide sequences for many influenza virus genes has begun to define the degree of sequence change required to produce outbreaks of epidemic influenza: such sequences also provide information concerning the origin and evolution of viruses and virus–virus interactions during replication.

We wish to thank our colleagues, Steve Inglis and Charles Penn, for helpful criticism, and Isabel Roditi and Jamie Robertson for providing nucleotide sequence data prior to publication. We are especially grateful to our typist, Mary Wright, for her accurate transcription of our manuscript. Our work in Cambridge is supported by grant PG979/140 from the Medical Research Council.

References

Abraham, G. (1979) *Virology* **97**, 177–182

Air, G. M. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7639–7643

Allen, H., McCauley, J., Waterfield, M. & Gething, M.-J. (1980) *Virology* **107**, 548–551

Baez, M., Taussig, R., Zazra, J. J., Young, J. F. & Palese, P. (1980) *Nucleic Acids Res.* **8**, 5845–5858
Baez, M., Zazra, J. J., Elliott, R. M., Young, J. F. & Palese, P. (1981) *Virology* **113**, 397–402
Barrett, T., Wolstenholme, A. J. & Mahy, B. W. J. (1979) *Virology* **98**, 211–225
Barry, R. D. & Mahy, B. W. J. (1979) *Br. Med. Bull.* **35**, 39–46
Barry, R. D., Ives, D. R. & Cruickshank, J. G. (1962) *Nature (London)* **194**, 1139–1140
Beaton, A. R. & Krug, R. M. (1981) *Nucleic Acids Res.* **9**, 4423–4436
Bentley, D. R. & Brownlee, G. G. (1982) *Nucleic Acids Res.* **10**, 5033–5042
Bishop, D. H. L., Huddleston, J. A. & Brownlee, G. G. (1982*a*) *Nucleic Acids Res.* **10**, 1335–1343
Bishop, D. H. L., Jones, K. L., Huddleston, J. A. & Brownlee, G. G. (1982*b*) *Virology* **120**, 481–489
Blaas, D., Patzelt, E. & Kuechler, E. (1982*a*) *Virology* **116**, 339–348
Blaas, D., Patzelt, E. & Kuechler, E. (1982*b*) *Nucleic Acids Res.* **10**, 4803–4812
Blok, J. & Air, G. M. (1982) *Virology* **121**, 211–229
Blok, J., Air, G. M., Laver, W. G., Ward, C. W., Lilley, G. G., Woods, E. F., Roxburgh, C. M. & Inglis, A. S. (1982) *Virology* **119**, 109–121
Bosch, F. X., Hay, A. J. & Skehel, J. J. (1978) in *Negative Strand Viruses and the Host Cell* (Mahy, B. W. J. & Barry, R. D., eds.), pp. 465–474, Academic Press, London
Both, G. W. & Sleigh, M. J. (1980) *Nucleic Acids Res.* **8**, 2561–2575
Both, G. W. & Sleigh, M. J. (1981) *J. Virol.* **39**, 663–672
Bouloy, M., Plotch, S. J. & Krug, R. M. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4886–4890
Bouloy, M., Morgan, M. A., Shatkin, A. J. & Krug, R. M. (1979) *J. Virol.* **32**, 895–904
Brand, C. M. & Skehel, J. J. (1972) *Nature (London) New Biol.* **238**, 145–147
Briedis, D. J. & Lamb, R. A. (1982) *J. Virol.* **42**, 186–193
Briedis, D. J., Conti, G., Munn, E. A. & Mahy, B. W. J. (1981) *Virology* **111**, 154–164
Briedis, D. J., Lamb, R. A. & Choppin, P. W. (1982) *Virology* **116**, 581–588
Caliguiri, L. A. & Gerstein, H. (1978) *Virology* **90**, 119–132
Carroll, A. R., McGeoch, D. J. & Mahy, B. W. J. (1975) *Colloq. Inst. Nat. Santé Rech. Med.* **47**, 95–102
Carter, M. J. & Mahy, B. W. J. (1982) *Arch. Virol.* **73**, 109–119
Caton, A. J. & Robertson, J. S. (1980) *Nucleic Acids Res.* **8**, 2591–2603
Compans, R. W., Klenk, H.-D., Caliguiri, L. A. & Choppin, P. W. (1970) *Virology* **42**, 880–889
Compans, R. W., Content, J. & Duesberg, P. H. (1972) *J. Virol.* **10**, 795–800
Davis, N. L. & Wertz, G. W. (1982) *J. Virol.* **41**, 821–832
Desselberger, U., Racianello, V. R., Zazra, J. J. & Palese, P. (1980) *Gene* **8**, 315–328
Dhar, R., Chanock, R. M. & Lai, C.-J. (1980) *Cell* **21**, 495–500
Duesberg, P. H. (1969) *J. Mol. Biol.* **42**, 485–499
Duesberg, P. H. (1970) in *The Biology of Large RNA Viruses* (Barry, R. D. & Mahy, B. W. J., eds.), pp. 301–323, Academic Press, London
Fang, R., Min Jou, W., Huylebroeck, D., Devos, R. & Fiers, W. (1981) *Cell* **25**, 315–323
Fields, S. & Winter, G. (1982) *Cell* **28**, 303–313
Fields, S., Winter, G. & Brownlee, G. G. (1981) *Nature (London)* **290**, 213–217
Follett, E. A. C., Pringle, C. R., Wunner, W. H. & Skehel, J. J. (1974) *J. Virol.* **13**, 394–399
Gething, M.-J., Bye, J., Skehel, J. J. & Waterfield, M. D. (1980) *Nature (London)* **287**, 301–306
Ghendon, Y.Z., Markushin, S. G., Klimov, A. I. & Hay, A. J. (1982) *J. Gen. Virol.* **63**, 103–111
Ghosh, K. & Ghosh, H. P. (1982) *Nucleic Acids Res.* **10**, 6341–6351
Glass, S. E., McGeoch, D. & Barry, R. D. (1975) *J. Virol.* **16**, 1435–1443
Griffith, I. P. (1975) in *Negative Strand Viruses* (Mahy, B. W. J. & Barry, R. D., eds.), pp. 121–132, Academic Press, London
Grist, N. R. (1979) *Br. Med. J.* **2**, 1632–1633
Guo, Y. J., Jin, F. G., Wang, P., Wang, M. & Zhu, J. M. (1983) *J. Gen. Virol.* **64**, 177–182
Hastie, N. D. & Mahy, B. W. J. (1973) *J. Virol.* **12**, 951–961
Hauptmann, R., Clarke, L. D., Mountford, R. C., Bachmayer, H. & Almond, J. W. (1983) *J. Gen. Virol.* **64**, 215–220
Hay, A. J. & Skehel, J. J. (1979) *J. Gen. Virol.* **44**, 599–608
Hay, A. J., Abraham, G., Skehel, J. J., Smith, J. C. & Fellner, P. (1977*a*) *Nucleic Acids Res.* **4**, 4197–4209
Hay, A. J., Lomniczi, B., Bellamy, A. R. & Skehel, J. J. (1977*b*) *Virology* **83**, 337–335
Hay, A. J., Skehel, J. J., Abraham, G., Smith, J. C. & Fellner, P. (1978) in *Negative Strand Viruses and the Host Cell* (Mahy, B. W. J. & Barry, R. D., eds.), pp. 313–323, Academic Press, London
Hay, A. J., Skehel, J. J. & McCauley, J. (1980) *Philos. Trans. R. Soc. London Ser. B* **288**, 341–348
Hay, A. J., Skehel, J. J. & McCauley, J. (1982) *Virology* **116**, 517–522
Heggeness, M. H., Smith, P. R., Ulmanen, I., Krug, R. M. & Choppin, P. W. (1982) *Virology* **118**, 466–470
Herz, C., Stavnezer, E., Krug, R. M. & Gurney, T. (1981) *Cell* **26**, 391–400
Hill, V. M., Marnell, M. & Summers, D. F. (1981) *Virology* **113**, 109–118
Hinshaw, V. S., Air, G. M., Gibbs, A. J., Graves, L., Prescott, B. & Karunak-Aran, D. (1982) *J. Virol.* **42**, 865–875
Hirst, G. K. (1962) *Cold Spring Harbor Symp. Quant. Biol.* **27**, 303–308
Hiti, A. L. & Nayak, D. P. (1982) *J. Virol.* **41**, 730–734
Hitti, A. L., Davis, A. R. & Nayak, D. P. (1981) *Virology* **111**, 113–124
Horisberger, M. A. (1980) *Virology* **107**, 302–305
Horisberger, M. A. (1982) *Virology* **120**, 279–286
Horne, R. W., Waterson, A. P., Wildy, P. & Farnham, A. E. (1960) *Virology* **11**, 79–98
Hoyle, L., Horne, R. W. & Waterson, A. P. (1961) *Virology* **13**, 448–459

Huddleston, J. A. & Brownlee, G. G. (1982) *Nucleic Acids Res.* **10**, 1029–1038
Inglis, S. C. & Brown, C. M. (1981) *Nucleic Acids Res.* **9**, 2727–2740
Inglis, S. C. & Mahy, B. W. J. (1979) *Virology* **95**, 154–164
Inglis, S. C., Carroll, A. R., Lamb, R. A. & Mahy, B. W. J. (1976) *Virology* **74**, 489–503
Inglis, S. C., McGeoch, D. J. & Mahy, B. W. J. (1977) *Virology* **78**, 522–536
Inglis, S. C., Barrett, T., Brown, C. M. & Almond, J. W. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3790–3794
Jackson, D. A., Caton, A. J., McCready, S. J. & Cook, P. R. (1982) *Nature (London)* **296**, 366–368
Kaptein, J. S. & Nayak, D. P. (1982) *J. Virol.* **42**, 55–63
Kawakami, K. & Ishihama, H. (1983) *J. Biochem (Tokyo)*, in the press
Klenk, H.-D. & Rott, R. (1980) *Curr. Top. Microbiol. Immunol.* **90**, 19–48
Krug, R. M. (1971) *Virology* **44**, 125–136
Krug, R. M. (1972) *Virology* **50**, 103–113
Krug, R. M. (1981) *Curr. Top. Microbiol. Immunol.* **93**, 125–149
Krug, R. M. & Etkind, P. R. (1973) *Virology* **56**, 334–348
Krug, R. M. & Soeiro, R. (1975) *Virology* **64**, 378–387
Krystal, M., Elliott, R. M., Benz, E. W., Jr., Young, J. F. & Palese, P. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4800–4804
Krystal, M., Buonagurio, D., Young, J. F. & Palese, P. (1983) *J. Virol.* **45**, 547–554
Lamb, R. A. & Choppin, P. W. (1976) *Virology* **74**, 504–519
Lamb, R. A. & Choppin, P. W. (1977) *J. Virol.* **23**, 816–819
Lamb, R. A. & Choppin, P. W. (1978) in *Negative Strand Viruses and the Host Cell* (Mahy, B. W. J. & Barry, R. D., eds.), pp. 229–237, Academic Press, London
Lamb, R. A. & Choppin, P. W. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4908–4912
Lamb, R. A. & Choppin, P. W. (1980) *Philos. Trans. R. Soc. London Ser. B* **288**, 327–333
Lamb, R. A. & Choppin, P. W. (1981) *Virology* **112**, 729–737
Lamb, R. A. & Lai, C.-J. (1980) *Cell* **21**, 475–485
Lamb, R. A. & Lai, C.-J. (1981) *Virology* **112**, 746–751
Lamb, R. A., Choppin, P. W., Chanock, R. M. & Lai, C.-J. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1857–1861
Lamb, R. A., Lai, C.-J. & Choppin, P. W. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4170–4174
Landsberger, F. R., Lenard, J., Paxton, J. & Compans, R. W. (1971) *Proc. Natl. Acad. Sci. U.S.A.* **68**, 2579–2583
Laver, W. G. (1978) *Virology* **86**, 78–87
Laver, W. G. & Valentine, R. C. (1969) *Virology* **38**, 105–119
Lubeck, M. D., Palese, P. & Schulman, J. L. (1979) *Virology* **95**, 269–274
Mahy, B. W. J. (1970) in *Biology of Large RNA Viruses* (Barry, R. D. & Mahy, B. W. J., eds.), pp. 392–415, Academic Press, London
Mahy, B. W. J. (1983) in *Influenza Virus Genetics* (Palese, P., ed.), Springer-Verlag, Vienna, in the press
Mahy, B. W. J. & Bromley, P. A. (1970) *J. Virol.* **6**, 259–268
Mahy, B. W. J., Hastie, N. D. & Armstrong, S. J. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1421–1424
Mahy, B. W. J., Hastie, N. D., Raper, R. H., Brownson, J. M. T. & Carroll, A. R. (1975) in *Negative Strand Viruses* (Mahy, B. W. J. & Barry, R. D., eds.), pp. 445–467, Academic Press, London
Mahy, B. W. J., Carroll, A. R., Brownson, J. M. T. & McGeoch, D. J. (1977) *Virology* **83**, 150–162
Mahy, B. W. J., Barrett, T., Briedis, D. J., Brownson, J. M. & Wolstenholme, A. J. (1980) *Philos. Trans. R. Soc. London Ser. B* **288**, 349–357
Mahy, B. W. J., Penn, C. R., Nichol, S. T., Briedis, D. J. & Barrett, T. (1981) in *Genetic Variation Among Influenza Viruses* (Nayak, D. P. & Fox, C. F., eds.), pp. 113–125, Academic Press, New York
Mark, G. E., Taylor J. M., Broni, B. & Krug, R. M. (1979) *J. Virol.* **29**, 744–752
Markoff, L. & Lai, C.-J. (1982) *Virology* **119**, 288–297
McCauley, J., Bye, J. Elder, K., Gething, M.-J., Skehel, J. J., Smith, A. & Waterfield, M. D. (1979) *FEBS Lett.* **108**, 422–426
McCauley, J. W., Mahy, B. W. J. & Inglis, S. C. (1982) *J. Gen. Virol.* **58**, 211–215
McGeoch, D. & Kitron, N. (1975) *J. Virol.* **15**, 686–695
McGeoch, D., Fellner, P. & Newton, C. (1976) *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3045–3049
Meier-Ewert, H. & Compans, R. W. (1974) *J. Virol.* **14**, 1083–1091
Min Jou, W., Verhoeyen, M., Devos, R., Saman, E., Fang, R., Huylebroeck, D., Fiers, W., Threlfall, G., Barber, C., Carey, N. & Emtage, S. (1980) *Cell* **19**, 683–696
Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472
Mowshowitz, S. L. (1981) in *The Replication of Negative Strand Viruses* (Bishop, D. H. L. & Compans, R. W., eds.), pp. 317–323, Elsevier/North-Holland, New York
Murti, K. G., Bean, W. J., Jr. & Webster, R. G. (1980) *Virology* **104**, 224–229
Nichol, S. T., Penn, C. R. & Mahy, B. W. J. (1981) *J. Gen. Virol.* **57**, 407–413
Palese, P. (1977) *Cell* **10**, 1–10
Penn, C. R., Blaas, D., Kuechler, E. & Mahy, B. W. J. (1982) *J. Gen. Virol.* **62**, 177–180
Pereira, M. S. (1982) in *Virus Persistence* (Mahy, B. W. J., Minson, A. C. & Darby, G. K., eds.), pp. 15–37, Cambridge University Press, Cambridge
Plotch, S. J. & Krug, R. M. (1977) *J. Virol.* **21**, 24–35
Ploch, S. J. & Krug, R. M. (1978) *J. Virol.* **25**, 579–586
Pons, M. W. (1971) *Virology* **46**, 149–160
Pons, M. W. (1976) *Virology* **49**, 789–792
Pons, M. W. & Rochovansky, O. M. (1979) *Virology* **97**, 183–189
Porter, A. G., Barber, C., Carey, N. H., Hallewell, R. A., Threlfall, G. & Emtage, J. S. (1979) *Nature (London)* **282**, 471–477
Porter, A. G., Smith, J. C. & Emtage, J. S. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5074–5078
Proudfoot, N. (1982) *Nature (London)* **298**, 516–517

Rees, P. J. & Dimmock, N. J. (1981) *J. Gen. Virol.* **53**, 125–132
Rees, P. J. & Dimmock, N. J. (1982) *J. Gen. Virol.* **59**, 403–408
Robertson, H. D., Dickson, E., Plotch, S. J. & Krug, R. M. (1980) *Nucleic Acids Res.* **8**, 925–942
Robertson, J. S. (1979) *Nucleic Acids Res.* **6**, 3745–3757
Robertson, J. S., Schubert, M. & Lazzarini, R. A. (1981) *J. Virol.* **38**, 157–163
Rochovansky, O. M. (1976) *Virology* **73**, 327–338
Rochovansky, O. M. (1981) in *The Replication of Negative Strand Viruses* (Bishop, D. H. L. & Compans, R. W., eds.), pp. 309–315, Elsevier/North-Holland, New York
Schild, G. C., Oxford, J. S. & Newman, R. W. (1979) *Virology* **93**, 569–573
Scholtissek, C. (1969) *Biochim. Biophys. Acta* **179**, 389–397
Scholtissek, C. (1978) *Curr. Top. Microbiol. Immunol.* **80**, 139–169
Scholtissek, C. (1979) *Adv. Genet.* **20**, 1–36
Scholtissek, C. & Rott, R. (1969) *J. Gen. Virol.* **4**, 125–137
Scholtissek, C. & von Hoyningen-Huene, V. (1980) *Virology* **102**, 13–20
Schubert, M., Keene, J. D., Herman, R. C. & Lazzarini, R. A. (1980) *J. Virol.* **34**, 550–559
Schulze, I. T. (1970) *Virology* **42**, 890–904
Schulze, I. T. (1972) *Virology* **47**, 181–196
Shaw, M. W., Lamb, R. A., Erickson, B. W., Briedis, D. J. & Choppin, P. W. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6817–6821
Sivasubramanian, N. & Nayak, D. P. (1982) *J. Virol.* **44**, 321–329
Skehel, J. J. (1972) *Virology* **49**, 23–36
Skehel, J. J. (1973) *Virology* **56**, 394–399
Skehel, J. J. & Burke, D. C. (1969) *J. Virol.* **3**, 429–438
Skehel, J. J. & Hay, A. J. (1978) *Nucleic Acids Res.* **5**, 1207–1219
Skehel, J. J., Douglas, A. R., Wilson, I. A. & Wiley, D. C. (1982) in *Virus Persistence* (Mahy, B. W. J., Minson, A. C. & Darby, G. K., eds.), pp. 215–226, Cambridge University Press, Cambridge
Sleigh, M. J., Both, G. W., Brownlee, G. G., Bender, V. J. & Moss, B. A. (1980) in *Structure and Variation in Influenza Virus* (Laver, G. & Air, G., eds.), pp. 69–80, Elsevier/North-Holland, New York
Smith, G. L. & Hay, A. J. (1982) *Virology* **118**, 96–108
Smith, W., Andrewes, C. H. & Laidlaw, P. P. (1933) *Lancet* **ii**, 66–68
Spooner, L. R. & Barry, R. D. (1977) *Nature* (*London*) **268**, 650–652
Staroff, A. & Bukrinskaya, A. G. (1981) *Arch. Virol.* **67**, 105–110
Stephenson, J. R., Hay, A. J. & Skehel, J. J. (1977) *J. Gen. Virol.* **36**, 237–248
Taylor, J. M., Hampson, A. W., Layton, J. E. & White, D. O. (1970) *Virology* **42**, 744–752
Thierry, F. & Danos, O. (1982) *Nucleic Acids Res.* **10**, 2925–2937
Ulmanen, I., Broni, B. A. & Krug, R. M. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7355–7359
Ulmanen, I., Broni, B. & Krug, R. M. (1983) *J. Virol.* **45**, 27–35
Van Rompuy, L., Min Jou, W., Huylebroeck, D. & Fiers, W. (1981) *Eur. J. Biochem.* **116**, 347–353
Van Rompuy, L., Min Jou, W., Huylebroeck, D. & Fiers, W. (1982) *J. Mol. Biol.* **161**, 1–11
Van Wyke, K. L., Hinshaw, V. S., Bean, W. J., Jr. & Webster, R. G. (1980) *J. Virol.* **35**, 24–30
Verhoeyen, M., Fang, R., Min Jou, W., Devos, R., Huylebroeck, D., Saman, E. & Fiers, W. (1980) *Nature* (*London*) **296**, 771–776
Ward, C. W. & Dopheide, T. A. (1980) in *Structure and Variation in Influenza Virus* (Laver, W. G. & Air, G. M., eds.), pp. 273–280, Elsevier/North-Holland, New York
Waterfield, M. D., Espelie, K., Elder, K. & Skehel, J. J. (1979) *Br. Med. Bull.* **35**, 57–63
Waterfield, M., Scrace, G. & Skehel, J. J. (1981) *Nature* (*London*) **289**, 422–424
Webster, R. G., Hinshaw, V. S., Bean, W. J., Van Wyke, K. L., Geraci, J. R., St. Aubin, D. J. & Petursson, G. (1981) *Virology* **113**, 712–724
Webster, R. G., Laver, W. G., Air, G. M. & Schild, G. C. (1982) *Nature* (*London*) **296**, 115–121
WHO Memorandum (1980) *Bull. W.H.O.* **58**, 585–591
Wiley, D. C., Wilson, I. A. & Skehel, J. J. (1981) *Nature* (*London*) **289**, 373–378
Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981) *Nature* (*London*) **289**, 366–373
Winter, G. & Fields, S. (1980) *Nucleic Acids Res.* **8**, 1965–1974
Winter, G. & Fields, S. (1981) *Virology* **114**, 423–428
Winter, G. & Fields, S. (1982) *Nucleic Acids Res.* **10**, 2135–2143
Winter, G., Fields, S. & Brownlee, G. G. (1981*a*) *Nature* (*London*) **292**, 72–75
Winter, G., Fields, S., Gait, M. J. & Brownlee, G. G. (1981*b*) *Nucleic Acids Res.* **9**, 237–245
Wolstenholme, A. J., Barrett, T., Nichol, S. T. & Mahy, B. W. J. (1980) *J. Virol.* **35**, 1–7
Wright, C. E. & Laver, W. G. (1978) *J. Mol. Biol.* **120**, 133–136
Wrigley, N. G. (1979) *Br. Med. Bull.* **35**, 35–38
Young, R. J. & Content, J. (1971) *Nature* (*London*) *New Biol.* **230**, 140–142