Ann. Rev. Biochem. 1987. 56:365–94

THE STRUCTURE AND FUNCTION OF THE HEMAGGLUTININ MEMBRANE GLYCOPROTEIN OF INFLUENZA VIRUS

Don C. Wiley

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

John J. Skehel

National Institute of Medical Research, Mill Hill, London NW7 1AA, England

CONTENTS

0066-4154/87/0701-0365$02.00

PERSPECTIVES AND SUMMARY

The determination of the three-dimensional structure of the hemagglutinin membrane glycoprotein, HA, of influenza virus in 1981 (1, 2) provided a structural interpretation for a large amount of data accumulated before then and sparked an intensification of research on the structure-function relationship of this glycoprotein. Besides an important role as a model membrane glycoprotein, the HA has been studied primarily because of its three major activities in the virus's infectious cycle. 1. The HA binds to a sialic-acid-containing receptor on a target cell to initiate the virus-cell interaction. 2. The HA mediates the entry of the virus into the cytoplasm by a membrane-fusion event. 3. The HA is the major surface antigen of the virus against which neutralizing antibodies are produced and as a consequence undergoes antigenic variation leading to recurrent epidemics of respiratory disease.

This review describes progress in understanding the molecular mechanisms of these processes. The location and properties of the receptor-binding site are described as well as an atomic model for virus-cell binding based on the X-ray structure of an HA-receptor fragment complex. A conformational transition is described, probably activated in vivo by the low pH in endosomes, which results in the activation of membrane-fusion activity concomitant with a molecular rearrangement that exposes a hydrophobic 'fusion' peptide. The proposed location of antibody-binding sites on the HA and of regions recognized as processed peptides by thymic lymphocytes are reviewed as well as how variation in the sequence of these regions results in recurrent epidemics.

Reviews of the crystal structure determination of the HA (3), the structure of the neuramidase glycoprotein of influenza virus (4), virus-cell receptors (5), and viral membrane fusion (6) have appeared as well as several compendia of articles on influenza viruses (7–11).

INFLUENZA VIRUS MEMBRANE

Influenza virus, like a number of other enveloped viruses, contains a lipid membrane that it obtains during maturation by budding from the plasma membrane of an infected cell (for review, see 12). The membrane contains cellular lipids, but the membrane proteins are coded by the virus. These proteins are inserted into the membrane during biosynthesis by the same 'signal peptide'–mediated process employed by cellular membrane and secretory proteins. Influenza virus has two membrane glycoproteins, the hemagglutinin (HA) and the neuraminidase (NA). Both of their three-dimensional structures have been determined by X-ray diffraction (1, 4).

THE STRUCTURE OF THE HEMAGGLUTININ GLYCOPROTEIN

The HA of the 1968 influenza virus strain A/Aichi/2/68(H_3N_2) is synthesized as a single polypeptide chain of 550 amino acids, which is subsequently cleaved by removal of arginine 329 into two chains, HA_1 (36,334 daltons) and HA_2 (25,750 daltons). These chains are covalently attached by a disulfide bond from HA_1 position 14 to HA_2 position 137 (13), and the two-chain monomers are associated noncovalently to form trimers on the surface of membranes (14). Bromelain treatment of virus yields a soluble trimer, BHA, of the extracellular region, containing all of HA_1 and the first 175 of 221 amino acids of HA_2, but missing the hydrophobic membrane-anchoring peptide (15). The three-dimensional structure of this trimer has been determined from X-ray studies (2, 12).

Three-Dimensional Structure

Figure 1 shows a schematic drawing of a monomer of the HA. The membrane end of the molecule is at the bottom of the drawing, and a vertical line shows the location of the threefold symmetry axis that would relate two more monomers in the trimer. Both the amino terminus of HA_1 and the C-terminus of BHA_2 are found at the extreme membrane end of the molecule. The structure is long, projecting 135 Å off the membrane. The molecule appears to have a globular domain on top of an elongated stem region. The globular region at the top contains only part of HA_1, while the stem contains parts of HA_1 and all of HA_2. The conformation of the chain is unusually extended. From the amino terminus at the membrane, the first 63 amino acids of HA_1 reach in a nearly extended structure 96 Å up the molecule before the first compact folding occurs. The globular region at the top contains an eight-strand antiparallel beta sheet. This domain contains the receptor-binding site (see RECEPTOR BINDING). The remainder of HA_1 returns to the stem, running nearly antiparallel to the initial stretch of HA_1.

The amino terminus of HA_2 is 22 Å from the C-terminus of HA_1, indicating that a conformational change accompanied the cleavage of the two polypeptide chains. The hydrophobic amino terminal peptide that is involved in membrane fusion (see VIRAL ENTRY AND MEMBRANE FUSION) is tucked into the trimer interface about 35 Å from the bottom of the molecule. The major structural feature of the stem is a hairpin loop of two alpha helixes (cylinders in Figure 1). The second helix is 80 Å long and forms the backbone of the stem region.

Independent local contacts stabilize the trimer in the globular and stem regions. The HA_1 globular domains make pairwise contacts near the upper

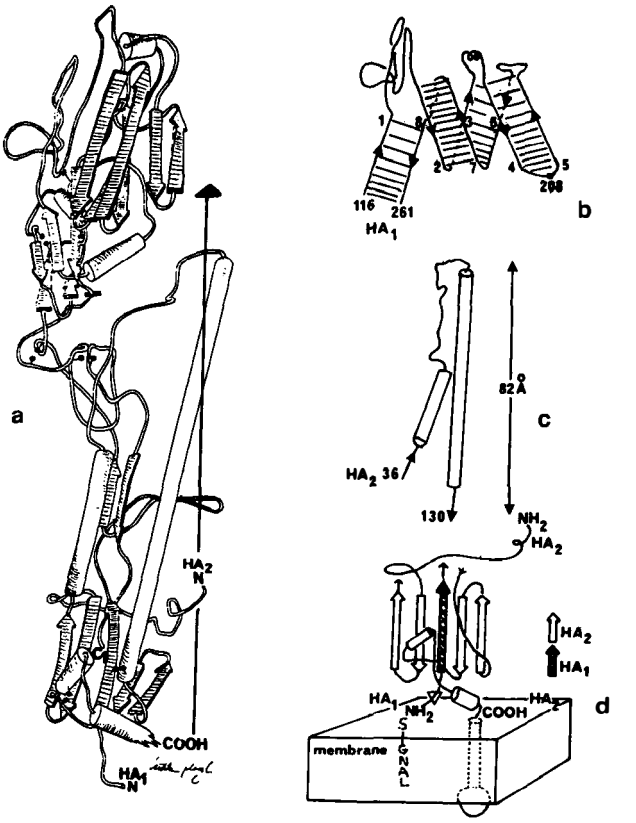


Figure 1 (*a*) Drawing of the three-dimensional structure of the influenza virus hemagglutinin. The terminal residues of HA_1 and HA_2 are labeled. The threefold symmetry axis relating molecules in the trimer is shown: the trimer is stabilized principally by packing of the long α-helixes (cylinders). Note that the amino terminus of HA_2, the 'fusion peptide,' tucks into the threefold contact. (*b*) The eight-stranded β-sheet structure and looped-out regions of the globular domain. (*c*) HA_2 residues 36–130 comprise an α-helical hairpin in the stem region of the monomer. Three of the long helixes, one from each monomer, pack together as a triple-stranded coiled-coil that stabilizes the trimer. (*d*) The membrane end of the molecule contains a five-stranded β-sheet. The central strand is the N terminus of HA_1, and the adjacent strands come from the C-terminal portions of the chain in HA_2. Broken lines suggest the path of the hydrophobic anchoring peptide removed by bromelain cleavage. The site of the oligosaccharide attached at HA_1 8 is shown as a triangle.

end of the molecule. A triple-stranded coiled-coil of alpha helixes formed from the top half of the long helix from each monomer provides the major contacts in the HA_2 region. These contacts are also stabilized by contacts between the second and third residues from each of the hydrophobic N-termini of HA_2 and a network of salt bridges in the lower part of the trimer.

Oligosaccharides

Six oligosaccharide chains are attached to asparagines of HA_1 (amino acid residues 8, 22, 38, 81, 165, 285), and one chain is attached to an asparagine of HA_2 (site number 154). All sites except 81 and 165 are processed, complex oligosaccharides (16). The most striking feature of the carbohydrate is its distribution (1). All sites except 165 are on the lateral surfaces of the molecule. One site, HA_1 8, is at the extreme membrane end of the molecule. The oligosaccharide at 165 appears to stabilize the oligomeric contacts between globular units at the top of the structure. The sequence of the oligosaccharides has been determined from one strain (17), however, no obligatory function has been assigned to them (1, 18, 19).

RECEPTOR BINDING

Sialic acid (*N*-acetyl neuraminic acid) is the only known essential component of cellular receptors for influenza type A virus. Oligosaccharides terminating in sialic acid are found on many cell-surface glycoproteins and glycolipids, and the removal of sialic acid from them by neuraminidase treatment destroys receptor binding, preventing infection (20). Furthermore, sialic-acid-containing glycolipids can serve as cellular receptors, restoring infectivity to neuraminidase-treated cells (21–23).

Receptor Site Structure

The influenza virus hemagglutinin's receptor-binding site is a pocket located on the distal end of the molecule (Figure 2*a*) and composed of amino acid side chains (tyr 98, his 183, glu 190, trp 153, leu 194) that are largely conserved in the numerous strains of the virus (1). Other conserved residues (cys 97, pro 99, cys 139, pro 147, tyr 195, arg 229) behind the pocket seem to stabilize the architecture of the site without being in a position to interact with the receptor. By contrast the perimeter of the surface of the pocket is composed of amino acid residues, which have varied during the antigenic drift that accompanies recurrent epidemics (2; Figure 6*b*).

The amino acid side chains forming the surface of the pocket, which are positioned such that they could make direct contact with a cellular receptor, are both polar and nonpolar (hydrophobic). Histidine 183, tyrosine 98, glutamic 190, and serine 136 have exposed side chain polar atoms that can act as hydrogen bond donors and acceptors. Leucine 194, tryptophan 153, and leucine 226 have exposed nonpolar side chains that have the potential to make van der Waal contacts with the hydrophobic portions of the receptor. In addition to these side chain atoms, sections of main chain from residues 225 to 228 and 131 to 137 have carbonyl oxygens and amide nitrogens of peptide bonds exposed with potential to interact with a receptor.

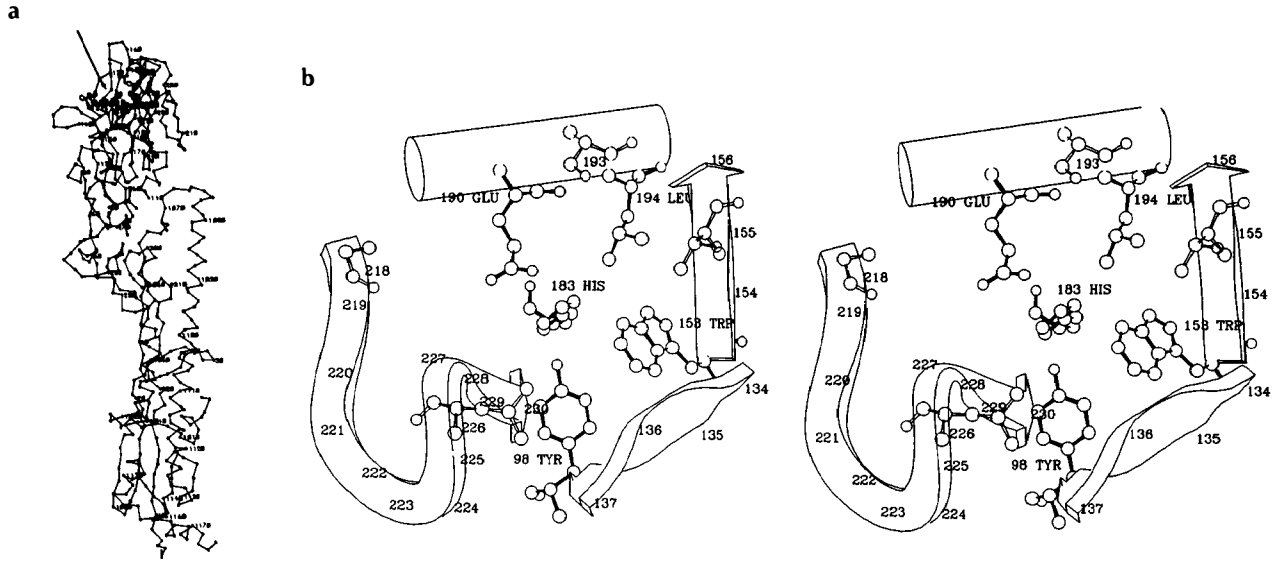


Figure 2 (*a*) Selected amino acids in and around the sialic acid receptor binding pocket are displayed on a schematic drawing of the alpha-carbon backbone of the hemagglutinin from the 1968 influenza virus X-31. (*b*) A detail of the same residues shown in a view of the top of the molecule in stereo. Leucine 226 projects into the pocket near the central residue Tyr 98. The site illustrated binds NeuAcα2→6Gal linkages, while glutamine at position 226 alters the site to bind preferentially NeuAcα2→3Gal linkages.

Topographically, the binding site is a depression, the bottom of which is formed by the phenolic hydroxyl of tyrosine 98 and the hydrophobic surface of tryptophan 153 (Figure 2*b*). The back wall of the site has histidine 183 projecting out to hydrogen bond with tyrosine 98 and both glutamic acid 190 and leucine 194 hanging down into the site from a short α-helix along the top of the back wall. Residues 224 to 228 form the 'left' edge and 133 to 137 the 'front' edge (Figure 2*b*). The right side of the site is an extended trough terminating at threonine 155.

In crystallographically refined electron density maps at 3 Å resolution (24), a number of water molecules are visible in the receptor-binding pocket that may mark positions where polar atoms of the receptor will bind [such patterns of bound water molecules being displaced by substrates have been observed in a number of cases in enzyme crystallography (e.g. see Ref. 25)].

Receptor Specificity Mutants

Confirmation that the pocket is the receptor-binding site was provided by the selection of single-amino-acid substitution mutants of the HA with altered receptor specificity. A comparison of the binding of different strains of influenza viruses of the H_3 subtype to erythrocytes derivatized with sialic acid in $\alpha 2\rightarrow 6$ and $\alpha 2\rightarrow 3$ linkages to galactose (NeuAc$\alpha 2\rightarrow$6Gal- or NeuAc$\alpha 2\rightarrow$3Gal-), had revealed distinct specificities (26). Preferential binding to sialic acid in $\alpha 2\rightarrow 6$ linkages, a trait of the 1968 HA, was found to correlate with a high sensitivity to neutralization of infection by glycoproteins (γ inhibitors) present in nonimmune horse serum (27). This allowed selection of receptor specificity variants by growth of virus in that serum (28). Five independently isolated mutants selected from two parental strains showed a reduction of affinity to $\alpha 2\rightarrow 6$ linkages and a marked increase in affinity for $\alpha 2\rightarrow 3$ linkages, so that four mutant viruses bind better to $\alpha 2\rightarrow 3$ than to $\alpha 2\rightarrow 6$ linkages, while one binds equally to both. In all five mutants, only a single-amino-acid substitution had occurred; in four cases glutamine was substituted for leucine at 226 and in one case methionine was substituted for leucine (28 and see Figure 2*b*), to produce the altered receptor specificity. Residue 226 is in the pocket described above (Figure 2*b*).

Structure of a Complex with Sialyl Lactose

At high concentration, the trisaccharide sialyl lactose inhibits the binding of influenza virus to erythrocytes (F. S. Escobar, D. C. Wiley, unpublished). To visualize directly the hemagglutinin's interaction with a receptor, .02 M $\alpha 2\rightarrow 6$ and $\alpha 2\rightarrow 3$ sialyl lactose were diffused into crystals of the wild-type 1968 HA (226 leu) and the specificity mutant (226 gln), respectively.

In both cases, difference Fourier electron density maps at 3 Å resolution show similar peaks in the receptor-binding site adjacent to position 226. Only

the structure of the specificity mutant, 226 gln, complexed with $\alpha 2 \rightarrow 3$ sialyl lactose, has been completed at this writing (W. Weis, J. C. Paulson, J. J. Skehel, D. C. Wiley, in preparation).

Model building and an exhaustive six-dimensional computer search of the fit of sialic acid to the difference electron density suggest two possible orientations, one of which is currently favored. This orientation places the glycerol side chain of sialic acid in the vicinity of tyrosine 98, glutamic 190, and serine 228. The methyl group of the N-acetyl moiety is located near tryptophan 153 and leucine 194, while the polar atoms of the amide link are near main chain polar atoms at residues 134 and 135. The carboxylate of sialic acid faces into the site near ser 136 and asn 137. The second possibility is in a similar position: C2 occupies about the same location, but the sugar ring is tilted and rotated 180° about an axis from C2 to C5 such that the glycerol and N-acetyl substituents switch locations and the carboxylate points up out of the site. Although the first alternative fits the density better, these models cannot be reliably distinguished at the current stage of the X-ray analysis.

No electron density is observed for the lactose moiety, indicating that it is spatially disordered, i.e. flexibly linked to the sialic acid.

Some alterations in the positions of side chains (< 1 Å) in the binding site are evident in the mutant's HA, which may be responsible for the altered specificity. Completing the X-ray structure analysis of the wild-type complex should help clarify this issue.

This description of binding to a receptor fragment now allows rationalization of some relative affinity measurements of various ligands and prepares the way for the design of inhibitors to block receptor binding and, therefore, infectivity of influenza virus.

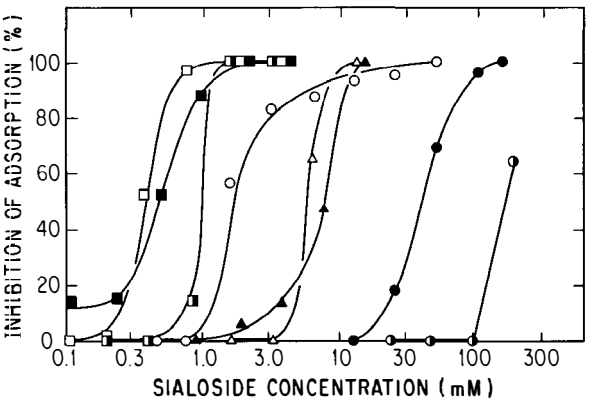
Ligand-Binding Comparisons

The binding of influenza virus to cell membranes is effectively irreversible, presumably due to the statistical cooperativity of the large number of hemagglutinin molecules covering the surface of the virus. To date there are no reports of measurements of the intrinsic site affinity of the hemagglutinin for its receptor. Recently, however, the relative binding affinities of a series of ligands have been measured by their ability to inhibit the binding of virus to sialidase-treated erythrocytes, which were derivatized with so few sialic acid residues that the statistical cooperativity of virus binding was markedly reduced (29). Four points are suggested by the data in Figure 3: 1. The α anomer of sialic acid (with the carboxylate axial) has over 10-fold higher affinity than the natural mixture of α and β, which is dominated (20 to 1) by β anomer. 2. A methyl group esterified to the carboxylate of sialic acid reduces affinity about fourfold. 3. The α anomer of the monosaccharide (α-methyl glycoside of NANA) binds almost as tightly (factor of 2 or 3) as an $\alpha 2 \rightarrow 6$-linked tetrasaccharide. 4. One hexasaccharide, doubly sialated, binds 10-fold

tighter than the monosaccharide. Points 2 and 3 are consistent with the structural model of ligand binding suggested above. Point 2: Because the carboxylate appears to face toward the protein surface, esterification with a methyl group would be expected to effect the interaction with the protein. Point 3: The X-ray results that two saccharides of sialic lactose are spatially disordered and only the sialic acid is tightly bound are consistent with the affinity measurements indicating that tri and tetrasaccharides have only slightly higher affinity than an α monosaccharide.

Affinity Mutants

Variations in receptor-binding affinity have been observed in natural strains of influenza viruses, and the importance of this in the epidemiology of the disease has been discussed (5, 30–32). Relative affinities have been inferred



Key:

Structure	Symbol
NeuAc	●
NeuAcβCH$_3$	◑
NeuAcαCH$_3$	○
NeuAcα2,3Galβ1,4GlcNAc - Lactose (LSTd)	△
NeuAcα2,3Galβ1,3GlcNAc - Lactose (LSTa)	▲
NeuAcα2,3Galβ1,3 > GlcNAc - Lactose (DSL) NeuAcα2,6	□
Galβ1,3 > GlcNAc - Lactose (LSTb) NeuAcα2,6	■
NeuAcα2,6Galβ1,4GlcNAc - Lactose (LSTc)	◨

Figure 3 Relative ligand affinities deduced from the inhibition of viral absorbtion to derivatized erythrocytes (from Ref. 29). See text.

by measuring the ability of virus to agglutinate a series of erythrocyte preparations whose receptors have been modified by partial sialidase or periodate treatment.

In one study, mutants of the 1934 virus, A/Puerto Rico/8/38, were selected with enhanced 'affinity' for receptors as inferred by the receptor gradient described above (33). Selection was accomplished by growth of virus in the presence of a mixture of antihemagglutinin monoclonal antibodies at subneutralizing concentrations. In the three mutants selected, a single amino acid substituted in the HA at positions 185, 231, and 244 correlated with enhanced receptor 'affinity.' In the first two cases the amino acids are in the second shell of residues around the receptor-binding site, while the third would appear to require a conformational alteration across the trimer interface adjacent to the receptor-binding site.

Twenty-two monoclonal-antibody-selected antigenic mutants of the 1968 X31 virus with single-amino-acid substitutions in their HA, when tested on a series of erythrocytes treated with increasing amounts of periodate, exhibited altered receptor binding (34). Some single-amino-acid substitutions at 156, 189, 193, and 135 increased the apparent 'affinity,' while others at 144, 145, 158, 188, 189 decreased 'affinity' as measured by this procedure. This demonstrates that single amino acids responsible for antigenic changes can also cause altered receptor-binding characteristics. Whether variation in receptor affinities is important in generating or limiting the recurrence of epidemics is not yet clear.

In a recent study of receptor binding, a series of mutant viruses selected using a single monoclonal antibody were shown to exhibit alteration in both receptor affinity and specificity as well as changes in the activation pH for membrane fusion activity by the HA (35) (See Table 1). Amino acid substitutions at 218 (G to R or E) and 226 (L to P) and remarkably a deletion of 7 amino acids from 224 to 230 cause reduction in the affinity for 2→6-linked sialic acid as measured by binding to derivatized erythrocytes, an increase in affinity for 2,3-linked sialic acid, and at the same time, substantial (.2 to .53 pH units) changes in the pH dependence of membrane fusion activity. Residue 218 and the 224–230 loop form part of the trimeric interface between the globular domains at the top of the HA.

These observations, which indicate that alterations at the trimer interface can also influence membrane fusion, demonstrate the structural potential for information to be transmitted between the binding site and the trimer interfaces and, therefore, possible interrelationships between the process of binding to cells and the triggering of entry by membrane fusion. No evidence indicates that cell binding influences the pH optimum of membrane fusion, but whether some ligand, natural or designed, could cause such a change during normal entry or as an antiviral device to cause a premature conformational change is unexplored.

Table 1 Receptor binding–membrane fusion mutants

Variant	Amino acid at position						Receptor specificity: Hemagglutination[a]		Receptor specificity: Absorption[b]		HI[c]	Change in fusion
	145	158	193	218	219	226	SAα2, 6Gal	SAα2, 3Gal	SAα2, 6Gal	SAα2, 3Gal	Eq α_2M	pH[d]
							HA titer		nmol/ml packed cells			
X-31	S	G	S	G	S	L	1024	0	13	>142	4096	—
63-E			N			P	512[t]	2[t]	30	93	0	+0.53
63-D						***	256[t]	256[t]	30	68	0	+0.20
63-3				E			1024	1024	10	54	1024	+0.43
v9A				R			256[t]	0	32	>142	1024	+0.37
v68x			R				1024	512	24	42	0	0
X-31/HS						Q	0	1924	53	46	0	0

[a] Hemagglutination assays were performed using 12.5% (v/v) human erythrocytes enzymatically modified to contain 37-nmol NeuAc/ml packed cells for the SAα2, 6Galβ1, 4GlcNAc (SAα2, 6Gal) sequences and 107-nmol NeuAc/ml packed cells for SAα2, 3Galβ1, 3GalNAc (SAα2, 3Gal) sequences. All of the viruses agglutinated native (untreated) erythrocytes but did not agglutinate sialidase-treated cells.

[b] These values represent the amount of sialic acid (in nmol/ml packed erythrocytes) required to bind 50% of the applied virus. The actual values are an average of the 50% binding point determined by assaying the amount of virus bound to derivatized cells and the amount of virus remaining in the supernatant.

[c] Hemagglutination inhibition (HI) by equine α_2 macroglobulin was determined at initial Eq α_2M concentrations of 3.5 mg/ml.

[d] A reasonance energy transfer assay was used to measure hemagglutinin fusion. By this assay X31 exhibited 50% fusion at pH 5.5.

[t] Indicates 'transient' agglutination where agglutination of derivatized erythrocytes was apparent after 30 min at room temperature, but no longer evident after 60 min.

***These variants have deleted amino acids 224–230 inclusive (RGLSSRI).

VIRAL ENTRY AND MEMBRANE FUSION

The two glycopolypeptides of hemagglutinin subunits, HA_1 and HA_2, are derived from the primary translation product of HA mRNA by proteolytic processing, which involves removal of the amino-terminal signal sequence and cleavage by an arginine-specific protease to generate the C-terminus of HA_1 and the N-terminus of HA_2 (36–40). The identities of the enzymes involved in the cleavage reaction are not known, but their actions are required for the production of hemagglutinins active in membrane fusion; uncleaved HAs have receptor-binding activity but are unable to fuse membranes (41–43) and as a consequence, viruses containing such molecules are not infectious (44, 45). These observations and their similarity to the required proteolytic processing of the precursor of Sendai virus fusion glycoprotein (46, 47) were the first indications of the involvement of HA in membrane fusion, and were reinforced by the findings that the N-terminal sequence of the Sendai virus glycopolypeptide generated by the cleavage reaction is analogous to the conserved, hydrophobic amino-terminal sequence of HA_2 (40, 48–50). Subsequently, experiments with SV40-HA recombinant virus-infected cells in which HA was expressed at the cell surfaces directly demonstrated its role in membrane fusion (51), a role that was also supported by demonstrations of the ability of purified hemagglutinins to mediate membrane fusion in vitro (52, 53).

Unlike viruses such as Sendai virus, which fuse with the surface membranes of cells, many membrane viruses appear to enter cells by a process involving endocytosis and to fuse their membranes with the membranes of the endosomes (6, 54, 55). Consequently, the observations that endosomal pH was about pH 5.0 instigated a series of experiments in which a number of viruses including influenza virus were shown to lyse erythrocytes, fuse liposomes, or fuse cells in culture between pH 5.0 and 6.5, depending on the particular virus (56–58).

A Low-pH-Induced Conformational Change Activates Fusion Activity

For influenza viruses, these findings prompted analyses of hemagglutinin structure at low pH in attempts to understand the molecular basis of fusion activation. Soluble hemagglutinins released from viruses by digestion with bromelain (BHA) were observed to form aggregates of sedimentation coefficient approximately 30S containing about eight hemagglutinins specifically at the pH of fusion. Incubation of BHA at this pH in the presence of ^{3}H-TX-100 indicated its acquisition of the ability to bind detergent, and the low-pH-specific association of BHA with liposomes was shown in experiments involving liposome-BHA complex flotation (59, 60).

In addition to these observations, which indicate the exposure of hydrophobic regions of the hemagglutinin in its low-pH conformation,

irreversible changes in structure were also shown by analyses of the proteolytic susceptibility of BHA, HA rosettes, and HA in virus particles after incubation at about pH 5.0 (59, 60). The resistance to proteolysis of native HA, clearly indicated by the quantitative isolation of BHA following extensive digestion of virus particles in high concentrations of bromelain (15), is in marked contrast to the susceptibility of HA in the low-pH conformation to digestion by trypsin, chymotrypsin, bromelain, proteinase K, and thermolysin (59–61). More detailed analysis of the tryptic digestion products shows that digestion in this case is restricted to the HA_1 glycopolypeptide yielding three characteristic fragments (Figure 4), one of which, HA_1 1–27, is linked to HA_2 by the single disulfide bond between HA_1 and HA_2 components of the subunit (HA_1 14S-S HA_2 137) and sediments in sucrose density gradients as a large aggregate. Thermolytic digestion of this aggregate converts it to a soluble product of sedimentation coefficient approximately 3S from which the N-terminal 23 residues of HA_2 have been removed (61). The conserved amino terminus of HA_2 appears, therefore, to be the hydrophobic region of the molecule involved in aggregate formation at low pH, its removal by thermolysin producing a soluble product. The other two tryptic peptides, HA_1 28–224 and HA_1 225–328, are recovered from the digestion products as a soluble complex of sedimentation coefficient about 2S, indicating that as a consequence of incubation at low pH, the membrane distal HA_1 domain is dissociated from the membrane proximal region of the molecule.

These observations were extended by the finding that HA_1 glycopolypeptides were quantitatively removed from viruses that had been incubated at pH 5 by cleavage of the disulfide bond between HA_1 and HA_2, which becomes susceptible to reduction specifically at low pH (62). Furthermore, the results of the sedimentation analysis and investigations of the subunit composition of the trypsin-solubilized HA_1 components by cross-linking with dimethyl suberimidate indicate that they are released as monomers and, therefore, that all of the interactions between individual HA_1 glycopolypeptides and other components of the hemagglutinin trimer are lost following incubation at low pH. Similar sedimentation analysis of hemagglutinin incubated at low pH indicates that at low concentrations in the presence of non-ionic detergents, the hemagglutinin completely dissociates (63, 64). At higher concentrations of hemagglutinin this appears not to be the case, and indeed cross-linking studies of the subunit composition of thermolysin-solubilized BHA_2 aggregates indicate that they remain trimeric (J. J. Skehel et al, unpublished).

The extensive nature of the changes in hemagglutinin conformation at low pH is also noted by electron microscopy, which appears to indicate that lengthening and dissociation of regions of the molecule occur (60, 65).

Changes in the antigenicity and immunogenicity of hemagglutinins resulting from incubation at low pH have also been reported and further indicate substantial changes in hemagglutinin conformation. The abilities of specific

Figure 4 (a) A schematic representation of the effect of *incubating* BHA at low pH. (*a*) pH 7.0 monomer structure. (*b*) After low pH, the hydrophobic N-terminal peptide of HA_2 is exposed (see text). (*c*) The low pH conformation of BHA can be cleaved in vitro by trypsin at HA_1 27, which leads to release of the 'top' globular domain from the stem, indicating a change in the tertiary interactions in the stem regions of the monomer. (*d*) The hydrophobic properties of the low pH form are lost after thermolytic removal of residues 1–23 of HA_2.

monoclonal antibodies to bind near intersubunit interfaces of hemagglutinins of the H_3 subtype were generally decreased, whereas the binding of antibodies recognizing sites A, E, and C (see *The Location of the Variable Antibody-Binding Sites*) was not affected, except for reactions with a number of antibodies specific for site C, which were enhanced (66–68). Studies with antibodies against A/PR/8/34 (H_1N_1) virus also indicated pH-dependent changes in hemagglutinin structure, although in these cases enhanced binding to sites A and E were the predominant effects (69).

In spectroscopic analyses, however, only small changes in fluorescence (52) and in near UV circular dichroism (CD) were noted (59). The results of far UV CD indicated that native HA and HA incubated at low pH were indistinguishable, implying that the changes in hemagglutinin conformation induced by low pH do not extend to large changes in secondary structure (59).

Overall, these analyses of the structure of hemagglutinin indicate that extensive rearrangement and loss of contact between components of the trimer that maintain their secondary structure occur as a consequence of incubation at low pH and are accompanied by or cause the extrusion of the hydrophobic amino terminus of HA_2 from its buried location in the trimer interface of the native molecule.

Membrane Fusion Mutants

Support for the proposed route of cell entry through endosomes has been obtained for a number of viruses by observing inhibitions of infections by reagents such as chloroquine, ammonium chloride, and amantadine, which elevate endosomal pH (eg. 70, 71). For influenza viruses in particular, high concentrations of amantadine have been shown to block cell entry (72, 73), and mutants that are resistant to this inhibition contain modifications specifically in their genes for hemagglutinin (74). The mutants retain membrane fusion activity, but their hemagglutinins assume the low pH conformation at higher pH than wild-type virus and mediate fusion at correspondingly higher pH. The HAs of the mutants have been sequenced to identify amino acid residues that influence the pH dependence of the change in conformation (74).

The structural locations of amino acid substitutions detected in the hemagglutinins of a collection of X-31 and Weybridge virus mutants are indicated in Figure 5*a*. The substitutions fall into two main groups on the basis of their proximity to the amino terminus of HA_2. Substitutions near the amino terminus involve amino acid residues that in the wild-type hemagglutinin stabilize the hydrophobic amino-terminal region in its buried position (Figure 5*b*). Residues that appear directly important in this context are HA_2 112-D and HA_1 17-H, which from hydrogen bonds with the terminal amino group and the amide nitrogens of residues 3, 4, and 5 of HA_2, and the carbonyl groups of residues 6 and 10 of HA_2, respectively. The amino acid substitutions at these

two locations listed in Table 2 involving either changes in side chain charge or length decrease the extent of hydrogen bonding and consequently the stability of this region of the molecule. Less directly, substitutions in neighboring residues such as HA_2 114 also destabilize the location of the amino terminus of HA_2 by introducing a positively charged amino group in close proximity to it. Amino acid substitutions in the HA_2 amino terminus itself include HA_2 6 I→M, HA_2 2 L→F, and HA_2 9 L→F, all of which maintain the uncharged

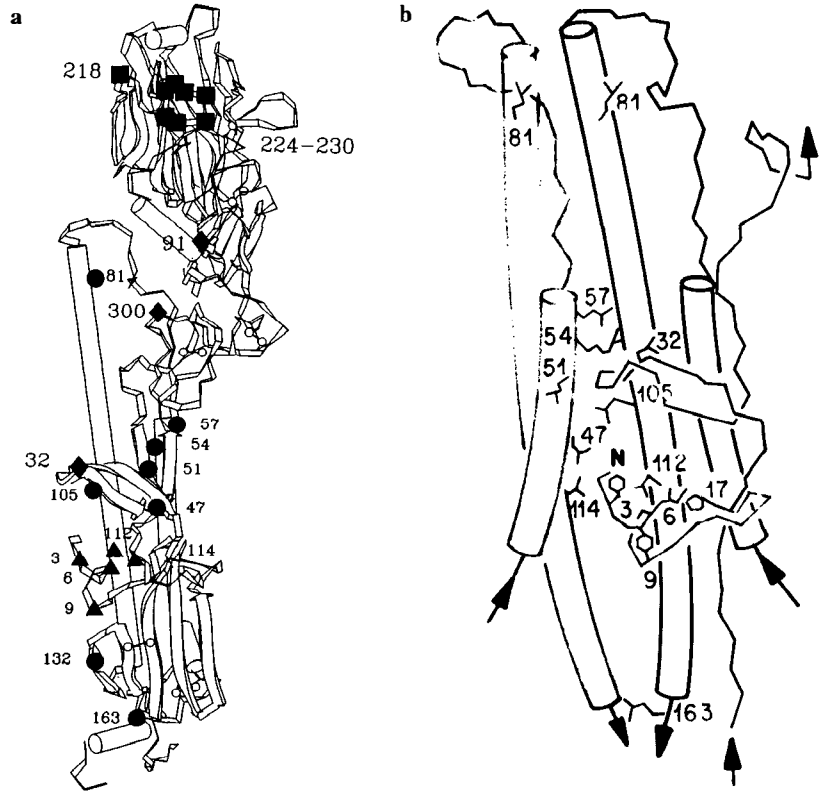


Figure 5 A schematic drawing of the HA monomer showing the location of single-amino-acid substitutions found in the membrane fusion mutants listed in Table 1. ■ indicates a substitution located in the HA_1-HA_1 interface, ● in the HA_2 trimer interface, ▲ in the vicinity of the N-terminal hydrophobic peptide of HA_2, and ◆ in the HA_1-HA_2 interface. Large number labels designate residues in HA_1. (*b*) An α carbon drawing of a detail from the HA trimer stem, showing the helical section of two HA_2 subunits (cylinders right and left) and a segment of HA_1 (thin line on right). The N-terminal peptide of HA_2, marked N, is tucked into the interface between two HA_2 subunits. The amino acid numbers show the positions of single-amino-acid substitutions in membrane fusion mutants.

Table 2 Amino acid substitutions in mutants that lyse erythrocytes at increased pH

	Residue		Amino acid substitution	ΔpH[b]
X-31				
1		HA_2 6	I→M	0.3
2	HA_1 207 +	HA_2 6	E→K; I→M	0.35
1a		HA_2 112	D→G	0.4
2a		HA_2 54	R→K	0.25
5a		HA_2 114	E→K	0.6
6a		HA_2 57+163	E→K; R→I	0.4
8a	HA_1 102 +	HA_2 6	V→M; I→M	0.5
4x		HA_2 81	E→G	0.3
6x		HA_2 9	F→L	0.4
12x		HA_2 112	D→N	0.8
aa1		HA_2 47	Q→R	0.35
aa2	HA_1 17		H→Q	0.9
aa4		HA_2 105	Q→K	0.3
aa9		HA_2 112	D→E	0.25
Weybridge[a]				
4		HA_2 69+78	E→G; Q→D	0.15
12	HA_1 91		R→Q	0.1
18		HA_2 81	I→S	0.1
2Y	HA_1 20+31		V→A; E→V	0.2
4Y		HA_2 112	D→G	0.4
A1	HA_1 113	HA_2 82	R→I; N→D	0.3
A2	HA_1 101	HA_2 43	K→E; S→L	0.1
A4	HA_1 32		R→G	0.2
A5	HA_1 91		R→L	0.3
A6	HA_1 300		R→S	0.3
A7	HA_1 324 +	HA_2 139	P→S; E→D	0.25
A8		HA_2 54	R→K	0.3
A9	HA_1 32 +	HA_2 82	R→G; N→Y	0.25
B1		HA_2 54+160	R→K; S→N	0.2
B5		HA_2 51+91	K→R; R→Q	0.45
B7		HA_2 114	E→K	0.5
B8		HA_2 67+102	D→N; M→R	0.1
B9	HA_1 221 +	HA_2 114	P→S; E→K	0.45
C5		HA_2 3	F→L	0.4
C6		HA_2 54	R→G	0.1
C7		HA_2 47	Q→L	0.45
C10		HA_2 54	R→S	0.1
C13		HA_2 51	K→N	0.2
C15	HA_1 205 +	HA_2 47	G→E; Q→K	0.3

[a] The residue numbers for the Weybridge mutants are related to the X-31 hemagglutinin sequence.
[b] ΔpH values are simple differences between the pHs at which hemolysis by the mutants and wild-type virus are 50% of maximum.

character of the region, but may influence its stability as a consequence of either changes in side chain size or orientation.

Mutant hemagglutinins in the second group contain substitutions in the interfaces between subunits in the trimer (Figure 5*b*). They all involve changes in charge and, therefore, influence local trimer stability by the loss of one or more salt bridges or hydrogen bonds. (See Table 2.) More importantly, they are located throughout the length of the molecule up to 100 Å distant one from another to indicate that different and widely separated residues influence the pH at which fusion activity is triggered. The regions included in this group have also been extended to HA_1 218 in the membrane distal HA_1-HA_1 interface by the observations that mutants selected with a specific anti-hemagglutinin monoclonal antibody that contain substitutions at this position also mediate fusion at higher pH than does wild-type virus (35). (See also Table 1.)

Mutants typical of both groups have also been isolated without the selective pressure of elevated endosomal pH. Viruses selected for the ability to replicate in MDCK cells without the addition of trypsin normally required to ensure precursor hemagglutinin cleavage were found to contain the substitution HA_1 17 H→R (75), and a naturally occurring mutation contained HA_2 132 D→N (76).

Finally, mutant genes for hemagglutinin have been constructed and expressed in SV40-HA recombinant virus-infected simian cells (77). All of the mutations reported were in the amino-terminal region of HA_2, involving residues HA_2 1, 4, and 11. Mutants containing the substitution HA_2 4 G→E mediated fusion at elevated pH, presumably as a consequence of destabilizing this region of the molecule in a similar fashion to the substitutions in the first group of mutants considered above. However, the ability of this mutant to mediate fusion abrogates an absolute requirement for conservation of the uncharged hydrophobic nature of this region, and in this regard is reminiscent of the HA_2 amino terminus in influenza C virus hemagglutinins, which contains aspartic acid residues at positions 5 and 6 (78, 79). Mutations at HA_2 1 G→E prevented membrane fusion and impaired lipid association, while the substitution HA_2 11 E→G had no apparent effect on membrane fusion per se but prevented polykaryon formation by cell-cell fusion. The molecular basis of these phenotypes is not clear at this stage, but such mutants may prove valuable in elucidating the mechanism of hemagglutinin-mediated membrane fusion.

The membrane fusion activities of a number of amantadine hydrochloride-selected mutants and of wild-type X-31 have also been investigated at various temperatures (53, 80). X-31, which at 37°C fuses membranes at pH 5.6, when incubated at 62°C fuses at pH 7.2. Mutants that at 37°C fuse at higher pH than the wild-type virus, fuse at lower temperatures, between 45°C and 60°C, at

pH 7.2. Analysis of the heat-induced changes in structure that accompany fusion activation indicates similar but more extensive changes than those observed following low-pH activation of fusion at physiological temperature. Overall, however, the observations support the interpretation that mutations that elevate the pH of fusion decrease the stability of the hemagglutinin and the notion of a requirement for substantial changes in hemagglutinin structure for fusion activation.

Summary of Conformational Changes

The three-dimensional structure of the low-pH-induced fusion-active conformation of the HA is not known. However the experiments discussed above provide a partial picture of that active conformation. 1. The hydrophobic N-terminal peptide of HA_2 is exposed as indicated by the hydrophobic properties and thermolytic sensitivity, and the loss of the hydrophobic properties after thermolytic treatment removes HA_2 1–23. 2. Extended segments of HA_1 in the stem region of the molecule are rearranged. The region near HA_1 27 becomes exposed to tryptic digestion and a disulfide at HA_1 14 to reduction. HA_1 can dissociate from HA_2 after one proteolytic nick or after reduction, probably indicating that the conformation of the extended HA_1 regions that contact HA_2 have rearranged to make less extensive contacts than in Figure 1. 3. The HA_1-HA_1 trimer interfaces between the top, globular domains of the molecule appear to be dissociated based on the exposure of a tryptic site at HA_1 224 and the recovery of the tryptic glycopolypeptide HA_1 28–328 as a monomer rather than a trimer. It remains possible that the HA_1-HA_1 interface is not dissociated but has rearranged and that the observation of monomers of HA_1 after trypsin treatment results from subsequent dissociation at the protein concentration used in the experiment. This seems less likely based on the observed antigenic changes in the HA_1-HA_1 interface region and the destabilizing nature of the mutants (Table 1) in the HA_1-HA_1 interface that affect the pH of membrane fusion. 4. A rearrangement of the HA_2-HA_2 trimer interface along its whole length from the virus membrane to the top of the long helix is suggested by the location of amino acid substitutions in fusion mutants (Figure 5*b*). Chemical cross-linking of the thermolysin-solubilized BHA_2 aggregate indicates that the stem region of the molecule remains trimeric, although the rearranged interface appears weakened since low-pH BHA can dissociate at sufficiently low protein concentrations.

Proton-induced conformational changes have been observed in the switching of oxy to deoxy hemoglobin and the assembly of tobacco mosaic virus (TMV) (81, 82). The current description of the HA conformational change suggests that certain regions of the molecule, the N-terminal peptide of HA_2 and regions of HA_1 in the stem of the molecule, may undergo not just a

change from one conformation to another, but may change from an ordered structure to a disordered ensemble of structures.

Membrane Fusion Mechanisms

The membrane fusion activity of the hemagglutinin has been measured both in vivo and in vitro and by a variety of techniques, including hemolysis (59, 83), polykaryon formation (84), resonance energy transfer (53, 85–87), liposome-cell fusion (88), spin-labeled phospholipid transfer (42), and electron microscopy (89, 90). In the main, determinations of the pH dependence of fusion by these procedures indicate correspondence between the pH of fusion and pH at which the characteristic changes in hemagglutinin structure occur. However, influenza-virus-mediated membrane fusion has also been reported to occur at neutral pH (86, 89, 90). In one case, liposomes containing negatively charged cardiolipin were used in the fusion assays, and the results obtained are probably not directly relevant to the hemagglutinin-mediated fusion process investigated using liposomes with more physiological lipid compositions (86, 87). The significance of other reports is, however, not clear. It may be that the efficiency of fusion in these systems is low, and simply results from a background level of fusion by viruses that fuse optimally at lower pH. Alternatively, the presence in virus preparations of small numbers of mutants able to fuse at elevated pH at 37°C may be involved, or it is possible that some ligands can induce the conformational change as discussed under *Affinity Mutants,* above. Nevertheless, the bulk of the evidence available at present indicates the requirement for a change in hemagglutinin conformation is triggered by low pH for influenza-virus-mediated membrane fusion.

What is the mechanism of this process? It is possible that the extruded HA_2 amino termini of neighboring molecules may associate to form hemagglutinin complexes active in membrane penetration or to expose areas of lipids in both virus and cellular membranes at which fusion may occur. However, the latter possibility seems inappropriate for fusion by viruses such as Sendai virus in which the protein involved in fusion is distinct from the receptor-binding protein and is therefore less likely as a general fusion mechanism. It is also possible in the absence of direct observations of membrane fusion in endosomes during influenza virus infections, that even though HA-mediated fusion can be demonstrated in vitro, the actual virus entry mechanism may result in membrane dissolution rather than in the formation of a continuous cellular-virus hybrid membrane. The favored scheme at present focuses on the conserved and hydrophobic nature of the amino terminus of HA_2 and suggests that as a result of the changes in hemagglutinin structure at low pH, direct interaction of this region of the molecule with an adjacent lipid bilayer is facilitated. Such an interaction may simply overcome the repulsive forces between the membranes, allowing their fusion, or may concomitantly destabilize the membrane and promote fusion as a consequence.

ANTIGENIC VARIATION

Antihemagglutinin antibodies neutralize virus infectivity, and as a consequence the antigenic variation for which influenza viruses are renowned involves extensive variation in the antigenic properties and structure of their hemagglutinin. For over 40 years the antigenicity of influenza isolates from outbreaks and epidemics has been monitored to ensure the inclusion in vaccines of variants closely related to the viruses in circulation (91), and more recently these comparative studies have been accompanied by nucleotide sequence analyses of the genes for the hemagglutinins of representative isolates (92–97).

The Location of the Variable Antibody-Binding Sites

The location of antibody-binding sites on the hemagglutinins of both subtypes of influenza A currently circulating (H_3 subtype viruses since 1968 and H_1 subtype since 1977) have been proposed based on the location of sequence variation in the HAs of natural isolates and on sites of single-amino-acid substitutions observed in antigenic mutants selected by growth in monoclonal antibodies (18, 30, 66, 98–103). Figure 6*a* shows the location on a schematic diagram of the HA of the naturally occurring sequence variation of the H_3 subtype HA during the last 18 years (five shaped symbols) and the single-site antigenic mutants selected from 1968 and 1972 parental strains with monoclonal antibodies (stars). (For a comparison with the H_1 subtype, see Ref. 103 and Figures 22 and 23 in Ref. 12). The variation is predominantly in the HA_1 glycopolypeptide, involving residues covering much of the surface of the distal domain of the molecule. The substitutions have been segregated into five regions designated A to E (2, 18), although some subdivision and overlap of these areas has been noted (66, 103).

It is currently thought that each of these regions is an antibody-binding area and that the HAs of new epidemic strains must include substitutions in each of these regions in order to escape the antisera produced during an earlier cycle of infection. The antigenic significance of the observed natural variation in regions A to E is strongly supported by the coincident mapping of monoclonal-antibody-selected variants into each of the regions (Figure 6*a*), because these variants are known to no longer bind the selecting antibody and as a consequence replicate efficiently in its presence. An additional indication that the natural variation is antigenically significant is the observation that the frequency of substitutions in HA_1, about .8% per year since 1968, is greater than that of silent mutation, .3%, indicating that it probably resulted from immune selective pressure.

The assumption that the location of a single-site substitution defines the site of antibody binding has been explored in two cases by crystallizing and determining the X-ray structure of the HAs from the monoclonal-antibody-

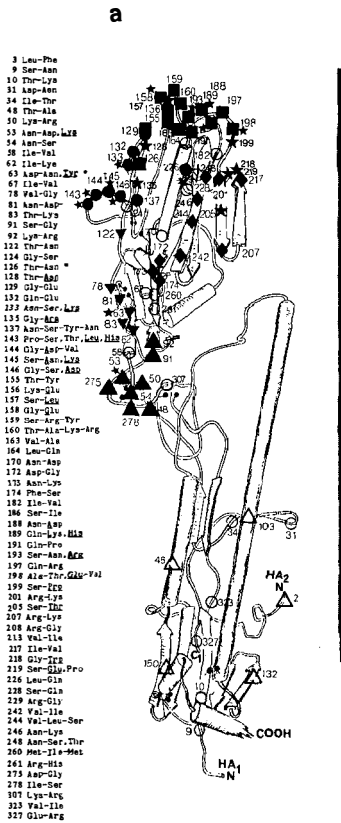


Figure 6 Natural variation since 1968 and monoclonal variants suggest the antibody-binding sites on the 1968 HA. (*a*) ●, Site A; ■, Site B; ▲, Site C; ◆, Site D; ▼, Site E. The symbols represent locations of natural sequence variation between 1968 and 1979. ★, single-site monoclonally selected variant, each star represents a separate variant; ★ plus site symbol, represents a site of natural variation that has also been observed in a monoclonally selected variant. Underlined amino acids in the list of amino acid substitutions were observed in monoclonally selected variants only, no underline indicates in natural variants only, first letter underlined indicates substitutions found in both natural and monoclonally selected variants. An asterisk indicates addition of an N glycosylation site, a minus indicates the loss of such a site. (*b*) Variable amino acid positions defining antibody-binding sites surround the conserved residues forming the receptor-binding pocket. The HA trimer is viewed from above, dotted spheres mark variable amino acid positions shown in Figure 6*a*. The center of the figure shows the α-helixes of HA_2 extending 'into' the page. The receptor sites are inside the crescent of variable residues. The remaining position of HA_1, which is not in the site or variable, is covered by the oligosaccharide position 165 (not shown).

selected mutants 146 G→D and 188 N→D (sites A and B) (104; W. Weis et al, unpublished). In both cases, the structural changes between the HA of the antigenic mutant and wild-type were confined to the immediate vicinity of the amino-acid substitution, demonstrating that for the antibody to recognize this mutation, it must bind directly to that region and illustrating that rather small, local variations in structure suffice to allow escape from neutralization by antibodies (104). Evidence consistent with the conclusion that the locations of amino-acid substitutions define the sites of antibody binding on the HA and that most substitutions influence antibody binding directly rather than from a distance, has also been obtained from electron microscopy of monoclonal antibody–hemagglutinin complexes using antibodies recognizing regions A, B, and E (105). No such direct information is presently available for binding sites C and D.

Because of the proximity of antibody-binding sites and the receptor-binding pocket, certain amino-acid substitutions influence both the binding of specific monoclonal antibodies and the specificity of receptor recognition (105a). As discussed above (RECEPTOR BINDING), a number of antihemagglutinin antibodies have also been observed to select mutants with different receptor-binding specificities which are, however, not antigenic mutants, since they continue to bind the selecting antibodies (33, 35). The sites of binding such antibodies are not known, but they do not in these cases appear to correspond with the locations of the amino acid sequence changes.

Some of the observed variation results in the addition or loss of an N-linked oligosaccharide (asterisk and dash in the list of Figure 6*a*). In the case of the monoclonal-antibody-selected mutant HA_1 63 D→N, the addition of the oligosaccharide has been proven to be the cause of the antigenic change, as mutant virus grown in the presence of tunicamycin and lacking the new oligosaccharide at 63 was not antigenically distinguishable by the antibody (18). A similar experiment on the natural epidemic strain A/VIC/3/75, which is glycosylated at 63, indicated that it would bind antibody produced against the 1968 strain (two epidemics earlier) when grown in tunicamycin and lacking the oligosaccharide at 63. Thus, carbohydrate, which is host-specific, can mask surfaces of the hemagglutinin from the immune system. Similar masking is seen in the area of HA_1 165 when the proposed antigenic sites of H_3 (glycosylated at 165) and H_1 (unglycosylated at 165) are compared (61, 103).

Figure 6*b* shows that the proposed antibody-binding sites surround the conserved receptor binding site on the HA. A similar observation has been made on the influenza NA (4) and has been proposed for Rhinovirus and Poliovirus based on their recent X-ray structure determinations (106, 107). Thus, even if antibodies could be directed at the conserved residues in the concave binding site, because of the large footprint of an antibody (108), they

would probably be dislodged by the variation that occurs readily on the rim of the binding site.

Recognition by Thymic Lymphocytes

Several studies of hemagglutinin immunogenicity and antigenicity have involved antibody production using synthetic peptides (109–113), and antibodies produced in this way have been of value in monitoring structural changes in the hemagglutinin and in studies of the molecular basis of antigenicity (114, 115). In the main, however, antipeptide antibodies have been inefficient in neutralizing virus infectivity and in binding to the native protein. This contrasts with the recognition of peptides by immune cells and the probable role of hemagglutinin fragments in stimulating cellular immunity.

Hemagglutinin-specific helper T-cell lines derived from peripheral blood lymphocytes from humans were found to be stimulated to proliferate by peptides HA_1 1–36, 105–140, 200–228, 306–328 of H_3 subtype hemagglutinin, and to a lesser extent, by peptides of similar sizes equivalent to all regions of HA_1. The most immunodominant peptide, HA_1 306–328, was recognized by both subtype-specific and cross-reactive clones isolated from these T-cell lines (116, 117).

Helper T-cells induced following infection of mice with A/Puerto Rico/8/34 (H_1N_1) were shown to recognize three regions of the HA, two of which were defined using peptides HA_1 109–120 and HA_1 290–310. The fine specificity of clones that recognized peptide HA_1 109–120 was analyzed using fragments of HA_1, various synthetic peptides from the region HA_1 109–120, a series of natural antigenic variants, and a monoclonal-antibody-selected variant to stimulate IL2 production (118–121). As a result, the importance for cellular recognition of residue HA_1 115, which is conserved in natural isolates of the H_1 subtype, was demonstrated. The third epitope defined in these studies involved recognition of HA_1 136.

Experiments with H_3 hemagglutinin-specific mouse cells, also involving the use of peptides, show the recognition of HA_1 48–68 and HA_1 128–148. In these studies, as with the studies of the H_1 subtype, clones of different specificity were differentiated by their abilities to recognize natural and monoclonal-antibody-selected antigenic variants, and also a fusion mutant (122, 123). The results suggest the importance of additional as yet undefined regions of HA_1 in helper T-cell recognition.

Overall hemagglutinin recognition by helper T-cells appears from these studies to be restricted to the HA_1 glycopolypeptide, including a number of regions that are recognized by antibodies. Peptides equivalent to these sites stimulate cells efficiently, which is consistent with the antigen processing roles proposed for antigen-presenting cells (124). In other studies, however, in which synthetic peptides (125) and isolated HA_1 and HA_2 glycopolypeptide

chains (126) were used to stimulate proliferation, helper T-cells that recognize components of HA_2 were detected and evidence was presented that they are more subtype cross-reactive than cells that recognize HA_1.

Cytotoxic T-cells appear to recognize both HA_1 and HA_2 components of the hemagglutinin. For H_2 subtype–specific cells, fragments of hemagglutinin have been reported to be ineffective in directing cell recognition (127). In other experiments, however, induction of cytotoxic cells was observed with a CNBr fragment, HA_2 103–123 (128), and this was confirmed using an equivalent synthetic peptide (129). In addition, a peptide analogue of HA_1 181–204 was shown to induce cytotoxic cells at about 10 times greater molar efficiency than the HA_2 peptide. Both of these peptides were recognized by subtype-specific cytotoxic cells (129). Experiments with a fusion protein containing the HA_2 glycopolypeptide produced in *E. coli* have also indicated the ability of this region of the hemagglutinin to generate target cells recognized by H_1 subtype–specific cytotoxic T-cells (130).

During many influenza infections, cytotoxic T-cells specific for HA represent only a minority of the immune cells induced, and as a consequence HA-specific cell clones are infrequently isolated. Following infection with A/Japan/305/57, however, about 45% of the clones obtained were found to be HA specific. Among these, about half were completely strain specific, reacting only with target cells expressing A/Japan/305/57 HAs, about one third were subtype specific, and the remainder either recognized hemagglutinins of both H_1 and H_2 subtypes or of all three subtypes tested, H_1, H_2, and H_3 (131).

Mechanisms of Antigenic Variation

Antigenic variation in influenza viruses involves two processes frequently referred to as antigenic shift and antigenic drift. The former is responsible for the introduction into the human population of viruses containing hemagglutinins antigenically similar to those of viruses circulating in animals of other species and birds. The most recent antigenic shift, which resulted in the Hong Kong influenza epidemic in 1968, was caused by a recombinant virus containing the gene for the H_3 hemagglutinin and the seven other genes in the virus-genome from an H_2N_2 virus circulating at that time in humans (132). Sequence analysis of the hemagglutinins of viruses of the H_3 subtype isolated from ducks and horses before 1968 strongly suggests that the recombination involved an avian influenza virus, and occurred presumably in a bird or human simultaneously infected with a human H_2N_2 virus and an avian virus of the H_3 subtype (133–136). The amino acid sequence homology between the Hong Kong (H_3) hemagglutinin and Asian influenza (H_2) hemagglutinins is approximately 36% HA_1 and 50% HA_2 (137, 138).

Antigenic variants of the H_3 subtype isolated between 1968 and 1986, however, vary in sequence homology by about 0.8% per year, and the process

involved in this variation is termed antigenic drift. Antigenic drift appears to involve the selection under antibody pressure of antigenic mutants with the ability to reinfect at least a proportion of the population. The precise mechanism of this selective process is unknown. The frequency at which antigenic mutants occur during selection with monoclonal antibodies is between 10^{-4} and 10^{-5} (139), and given that antibodies against any of the five antigenic regions neutralize virus infectivity, antigenic drift to an extent that would allow reinfection of the majority of the population would occur spontaneously at a very low frequency. Changes at each of the antigenic sites may, on the other hand, occur during reinfections by antigenic mutants changed in only one or two sites in individuals who develop only partial immunity during initial infections. Analysis of the variety of antibody specificities in post-infection human sera and of the restricted ability of human sera to neutralize monoclonal-antibody-selected mutants suggest that this is the case (140, 141). Information on the frequency of reinfections from more extensive serological surveys is required for assessment of this possibility.

CONCLUSION

The molecular models emerging for the major membrane activities of the HA glycoprotein provide a challenge for the future. Can measurements of intrinsic site affinities coupled with structural analysis of HA-receptor analogue complexes provide a description of the events and energetics of virus-to-cell binding? Can the binding potential of the HA receptor-binding site be sufficiently understood to allow binding inhibitors to be designed that could prevent infection? What is the mechanism of membrane fusion? Can a description of the conformational change required for membrane fusion activity be used as the basis for designing strategies to prevent the change, or to trigger it to occur prematurely, thus preventing membrane fusion required for infection?

Using current information and reagents—site-defined monoclonal antibodies, single-site antigenic mutants, extensive sequence information for the natural antigenic variants, and collections of post-infection antisera from a decade of recurrent influenza—can the spectrum of antihemagglutinin antibodies produced in individuals and their consequences be determined to a degree that allows a detailed path for the selection of new epidemic strains to be seen? What is a plausible explanation for the generation of recurrent epidemics?

ACKNOWLEDGMENTS

DCW acknowledges support from NIH AI-13654 and NSF DMB 85-02920.

Literature Cited

1. Wilson, I. A., Skehel, J. J., Wiley, D. C. 1981. *Nature* 289:368–73
2. Wiley, D. C., Wilson, I. A., Skehel, J. J. 1981. *Nature* 289:373–78
3. Wiley, D. C., Wilson, I. A., Skehel, J. J. 1984. In *Biological Macromolecules and Assemblies,* Vol. I: *Viral Structures,* ed. F. Jurnak, A. McPherson, pp. 299–336. New York: Wiley
4. Colman, P. M., Varghese, J. N., Laver, W. G. 1983. *Nature* 303:41–47
5. Paulson, J. C. 1985. In *The Receptors,* Vol. II, pp. 131–219. New York: Academic
6. White, J., Kielian, M., Helenius, A. 1983. *Q. Rev. Biophys.* 16:151–95
7. Laver, W. G., Chu, C. M., eds. 1985. *The Origin of Pandemic Influenza Viruses.* New York: Elsevier-North Holland
8. Nayak, D. P., ed. 1981. *ICN-UCLA Symp. Mol. Cell. Biol.* Vol. 21. New York: Academic
9. Palese, P., Kingsbury, D. W., eds. 1983. *Genetics of Influenza Viruses.* New York: Springer-Verlag
10. Laver, W. G., Air, G., eds. 1985. *Immune Recognition of Protein Antigens. In Current Communication in Molecular Biology.* New York: Cold Spring Harbor Lab.
11. Kendal, A. P., Patriarca, P. A., eds. 1985. *Options for the Control of Influenza Virus. UCLA Symp. Mol. Cell. Biol., Vol. 36.* New York: Liss
12. Wiley, D. C. 1985. In *Virology,* ed. B. N. Fields, pp. 45–68. New York: Raven
13. Ward, C. W. 1981. *Curr. Top. Microbiol. Immunol.* 94/95:1–17
14. Wiley, D. C., Skehel, J. J., Waterfield, M. 1977. *Virology* 79:446–48
15. Brand, C. M., Skehel, J. J. 1972. *Nature* 238:145–47
16. Ward, C. W., Dopheide, T. A. 1980. *Virology* 103:37–53
17. Keil, W., Geyer, R., Dabrowski, J., Dabrowski, U., Niemann, H., et al. 1985. *EMBO J.* 4:2711–20
18. Skehel, J. J., Stevens, D. J., Daniels, R. S., Douglas, A. R., Knossow, M., et al. 1984. *Proc. Natl. Acad. Sci. USA* 81:1779–83
19. Raymond, F. L., Caton, A. J., Cox, N. J., Kendal, A. P., Brownlee, G. G. 1983. *Nucleic Acids Res.* 11:7191–203
20. Gottschalk, A. 1959. In *The Viruses,* ed. F. V. M. Bumet, W. V. M. Stanley, 3:51–61. New York: Academic
21. Bergelson, L. D., Bukrinskaya, A. G., Prokazova, N. V., Shaposhnikova, G. I., Kocharov, S. L., et al. 1982. *Eur. J. Biochem.* 128:467–74
22. Suzuki, Y., Matsunaga, M., Matsumoto, M. 1985. *J. Biol. Chem.* 260:1362–65
23. Bukrinskaya, A. G., Kornilaeva, G. V., Vorkunova, N. K., Timofeeva, N. G., Shaposhnikova, G. I., Bergelson, L. G. 1982. *Vopr. Virusol.* 1982:661–66
24. Knossow, M., Lewis, M., Rees, D., Wilson, I. A., Wiley, D. C. 1986. *Acta Crystallogr.* In press
25. James, M. N. G., Sielecki, A. R. 1984. *Biochemistry* 24:3701–13
26. Rogers, G. N., Paulson, J. C. 1982. *Fed. Proc.* 41:5880
27. Rogers, G. N., Pritchett, T., Paulson, J. C. 1983. *Fed. Proc.* 42:2181
28. Rogers, G. N., Paulson, J. C., Daniels, R. S., Skehel, J. J., Wilson, I. A., Wiley, D. C. 1983. *Nature* 304:76–78
29. Pritchett, T., Brosmer, R., Paulson, J. C. 1986. Submitted for publication
30. Fazekas de St. Groth, S. 1978. In *Topics in Infectious Diseases,* ed. W. G. Laver, H. Bachmayer, R. Weil, 3:25–48. Vienna: Springer-Verlag
31. Underwood, P. A. 1982. *J. Gen. Virol.* 62:153–69
32. Underwood, P. A. 1985. *Arch. Virol.* 84:53–62
33. Yewdell, J. W., Caton, A. J., Gerhard, W. 1986. *J. Virol.* 57:623–28
34. Underwood, A., Skehel, J. J., Wiley, D. C. 1987. *J. Virol.* 61(1):206–8
35. Daniels, R. S., Jeffries, S. A., Yates, P. A., Schild, G. C., Rogers, G. N., et al. 1986. Submitted for publication
36. Elder, K. T., Bye, J. M., Skehel, J. J., Waterfield, M. D., Smith, A. E. 1979. *Virology* 95:343–50
37. Air, G. M. 1979. *Virology* 97:468–72
38. McCauley, J., Bye, J. M., Elder, K. T., Gething, M. J., Skehel, J. J., et al. 1979. *FEBS Lett.* 108:422–26
39. Garten, W., Bosch, F. X., Linder, D., Rott, R., Klenk, H.-D. 1981. *Virology* 115:361–74
40. Skehel, J. J., Waterfield, M. D. 1975. *Proc. Natl. Acad. Sci. USA* 72:93–97
41. Huang, R. T. C., Rott, R., Klenk, H.-D. 1981. *Virology* 110:243–47
42. Maeda, T., Kawasaki, K., Ohnishi, S. 1981. *Proc. Natl. Acad. Sci. USA* 78:4133–37
43. White, J., Matlin, K., Helenius, A. 1981. *J. Cell Biol.* 89:674–79
44. Klenk, H.-D., Rott, R., Orlich, M., Blodorn, J. 1975. *Virology* 68:426–39

45. Lazarowitz, S. G., Choppin, P. W. 1975. *Virology* 68:440–54
46. Homma, M., Ouchi, M. 1973. *J. Virol.* 12:1457–65
47. Scheid, A., Choppin, P. W. 1974. *Virology* 57:475–90
48. Gething, M. J., White, J., Waterfield, M. 1978. *Proc. Natl. Acad. Sci. USA* 75:2737–40
49. Scheid, A., Graves, M., Silver, S., Choppin, P. W. 1978. In *Negative Strand Viruses and the Host Cell,* ed. B. W. J. Mahy, R. D. Barry, pp. 181–93. London: Academic
50. Richardson, C. D., Scheid, A., Choppin, P. W. 1980. *Virology* 105:205–22
51. White, J., Helenius, A., Gething, M. J. 1982. *Nature* 300:658–59
52. Sato, S. B., Kawasaki, K., Ohnishi, S. I. 1983. *Proc. Natl. Acad. Sci. USA* 80:3153–57
53. Wharton, S. A., Skehel, J. J., Wiley, D. C. 1986. *Virology* 149:27–35
54. Simons, K., Garoff, H., Helenius, A. 1982. *Sci. Am.* 246:58–66
55. Helenius, A., Marsh, M., White, J. 1980. *Trends Biochem. Sci.* 5:104–6
56. Baananen, T., Kaariainen, L. 1979. *J. Gen. Virol.* 43:593–601
57. Matlin, K., Reggio, H., Simons, K., Helenius, A. 1982. *J. Mol. Biol.* 156:609–31
58. Lenard, J., Miller, D. K. 1981. *Virology* 110:479–82
59. Skehel, J. J., Bayley, P. M., Brown, E. B., Martin, S. R., Waterfield, M. D., et al. 1982. *Proc. Natl. Acad. Sci. USA* 79:968–72
60. Doms, R. W., Helenius, A., White, J. 1985. *J. Biol. Chem.* 260:2973–81
61. Daniels, R. S., Douglas, A. R., Skehel, J. J., Waterfield, M. D., Wilson, I. A., Wiley, D. C. 1985. See Ref. 7, pp. 1–7
62. Graves, P. N., Schulman, J. L., Young, J. F., Palese, P. 1983. *Virology* 126:106–16
63. Nestorowicz, A., Laver, W. G., Jackson, D. C. 1985. *J. Gen. Virol.* 66:1687–95
64. Doms, R. W., Agnew, W., Helenius, A. 1986. Submitted for publication
65. Ruigrok, R. W. H., Wrigley, N. G., Calder, L. J., Cusack, S., Wharton, S. A., et al. 1986. *EMBO J.* 5:41–49
66. Daniels, R. S., Douglas, A. R., Skehel, J. J., Wiley, D. C. 1983. *J. Gen. Virol.* 64:1657–62
67. Webster, R. G., Brown, L. E., Jackson, D. C. 1983. *Virology* 126:587–99
68. Jackson, D. C., Nestorowicz, A. 1985. *Virology* 145:72–83
69. Yewdell, J. W., Gerhard, W., Bachi, T. 1983. *J. Virol.* 48:239–48
70. Helenius, A., Marsh, M., White, J. 1982. *J. Gen. Virol.* 58:47–61
71. Miller, D. K., Lennard, J. 1981. *Proc. Natl. Acad. Sci. USA* 78:3605–9
72. Skehel, J. J., Hay, A. J., Armstrong, J. A. 1977. *J. Gen. Virol.* 38:97–110
73. Kato, N., Eggers, H. J. 1969. *Virology* 37:632–41
74. Daniels, R. S., Downie, J. C., Hay, A. J., Knossow, M., Skehel, J. J., et al. 1985. *Cell* 40:431–39
75. Rott, R., Orlich, M., Klenk, H.-D., Wang, M. L., Skehel, J. J., Wiley, D. C. 1985. *EMBO J.* 3:3329–32
76. Doms, R. W., Gething, M. J., Henneberry, J., White, J., Helenius, A. J. 1986. *Virology* 57:603–13
77. Gething, M. J., Doms, R. W., York, D., White, J. 1986. *J. Cell Biol.* 102:11–23
78. Nakada, S., Treager, R. S., Krystal, M., Aaronson, R. P., Palese, P. 1984. *J. Virol.* 50:118–24
79. Pfeifer, J. B., Compans, R. W. 1984. *Virus Res.* 1:281–96
80. Ruigrok, R. W. H., Martin, S. R., Wharton, S. A., Skehel, J. J., Bayley, P. M., Wiley, D. C. 1986. *Virology.* 155:484–97
81. Perutz, M. F., Fermi, G., Shih, T.-B. 1984. *Proc. Natl. Acad. Sci. USA* 81:4781–84
82. Butler, P. J. G., Durham, A. C. H., Klug, A. 1972. *J. Mol. Biol.* 72:1–18
83. Maeda, T., Ohnishi, S. I. 1980. *FEBS Lett.* 122:283–87
84. White, J., Helenius, A., Kartenbeck, J. 1983. *EMBO J.* 1:217–22
85. Wilschut, J., Hoekstra, D. 1984. *Trends Biochem. Sci.* 9:479–83
86. Stegmann, T., Hoekstra, D., Scherphof, G., Wilschut, J. 1985. *Biochemistry* 24:3107–33
87. Stegmann, T., Hoekstra, D., Scherphof, G., Wilschut, J. 1986. *J. Biol. Chem.* In press
88. Van Meer, G., Bavoust, J., Simons, K. 1985. *Biochemistry* 24:3593–602
89. Huang, R. T. C., Wahn, K., Klenk, H.-D., Rott, R. 1980. *Virology* 104:294–302
90. Haywood, A. M., Boyer, P. P. 1985. *Proc. Natl. Acad. Sci. USA* 82:4611–15
91. Pereira, M. S. 1982. In *Virus Persistence, Symp. 33, Soc. Gen. Microbiol.,* eds. B. W. J. Mahy, A. C. Minson, G. K. Darby, pp. 15–37. Cambridge Univ. Press
92. Sleigh, M. J., Both, G. W., Un-

derwood, P. A., Bender, V. J. 1981. *J. Virol.* 37:845–53
93. Both, G. W., Sleigh, M. J. 1981. *J. Virol.* 39:663–72
94. Both, G. W., Sleigh, M. J., Cox, N. J., Kendal, A. P. 1983. *J. Virol.* 48:52–60
95. Skehel, J. J., Daniels, R. S., Douglas, A. R., Wiley, D. C. 1983. *Bull. WHO* 61:671–76
96. Daniels, R. S., Douglas, A. R., Skehel, J. J., Wiley, D. C. 1985. *Bull. WHO* 63:273–77
97. Laver, W. G., Air, G. M., Dopheide, T. A., Ward, C. W. 1980. *Nature* 283:454–57
98. Gerhard, W., Webster, R. G. 1978. *J. Exp. Med.* 148:383–92
99. Moss, B. A., Underwood, P. A., Bender, V. J., Whittaker, R. G. 1980. In *Structure and Variations in Influenza Virus,* ed. W. G. Laver, G. M. Air, pp. 329–38. Amsterdam: Elsevier
100. Laver, W. G., Air, G. M., Webster, R. G., Gerhard, W., Ward, C. W., Dopheide, T. A. 1979. *Virology* 98:226–37
101. Gerhard, W., Yewdell, J., Frankel, M. E., Webster, R. 1981. *Nature* 290:713–17
102. Newton, S. E., Air, G. M., Webster, R. G., Laver, W. G. 1983. *Virology* 128:495–501
103. Caton, A. J., Brownlee, G. G., Yewdell, J. W., Gerhard, W. 1982. *Cell* 31:417–27
104. Knossow, M., Daniels, R. S., Douglas, A. R., Skehel, J. J., Wiley, D. C. 1984. *Nature* 311:678–80
105. Wrigley, N. G., Brown, E. B., Daniels, R. S., Douglas, A. R., Skehel, J. J., Wiley, D. C. 1983. *Virology* 131:308–14
105a. Daniels, R. S., Douglas, A. R., Skehel, J. J., Wiley, D. C., Naeve, C. W., et al. 1984. *Virology* 138: 174–77
106. Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., et al. 1986. *Nature* 317:145–53
107. Hogle, J. M., Chow, M., Filman, D. J. 1986. *Science* 229:1358–65
108. Amit, A. G., Mariuzza, R. A., Phillips, S. E. V., Poljak, R. J. 1986. *Science* 233:747–53
109. Green, N., Alexander, H., Olson, A., Alexander, S., Shinnick, T. M., et al. 1982. *Cell* 28:477–87
110. Muller, G. M., Shapira, M., Arnon, R. 1982. *Proc. Natl. Acad. Sci. USA* 79:569–73
111. Jackson, D. C., Murray, J. M., White, D. O., Fagan, C. N., Tregear, G. W. 1982. *Virology* 120:273–76
112. Shapira, M., Jibson, M., Muller, G., Arnon, R. 1984. *Proc. Natl. Acad. Sci. USA* 81:2461–65
113. Atassi, M. Z., Webster, R. G. 1983. *Proc. Natl. Acad. Sci. USA* 80:840–44
114. White, J., Wilson, I. 1986. Submitted for publication
115. Wilson, I. A., Niman, H. L., Houghten, R. A., Cherenson, A. R., Connolly, M. L., Lerner, R. A. 1984. *Cell* 37:767–78
116. Lamb, J. R., Eckels, D. D., Lake, P., Woody, J. N., Green, N. 1982. *Nature* 300:66–69
117. Lamb, J. R., Green, N. 1983. *Immunology* 50:659–66
118. Hackett, C. J., Dietzschold, B., Gerhard, W., Ghrist, B., Knorr, R., et al. 1983. *J. Exp. Med.* 158:294–302
119. Hurwitz, J. L., Haber-Katz, E., Hackett, C. J., Gerhard, W. 1984. *J. Immunol.* 133:3371–77
120. Gerhard, W., Hackett, C., Melchers, F. J. 1983. *J. Immunol.* 130:2379–85
121. Hackett, C. J., Hurwitz, J. L., Dietzschold, B., Gerhard, W. U. 1985. *Immunology* 135:1391–94
122. Mills, K. H. G., Skehel, J. J., Thomas, T. B. 1986. *J. Exp. Med.* 163:1477–90
123. Mills, K. H. G., Skehel, J. J., Thomas, T. B. 1986. *Eur. J. Immunol.* In press
124. Unanue, E. R. 1984. *Ann. Rev. Immunol.* 2:395–428
125. Attasi, M. Z., Kurisaki, J. I. 1984. *Immunol. Commun.* 13:539–51
126. Katz, J. M., Laver, W. G., White, D. O., Anders, D. M. 1985. *J. Immunol.* 134:616–22
127. Morrison, L. A., Lukacher, A. E., Braciale, V. L., Fan, D. P., Braciale, T. J. 1986. *J. Exp. Med.* 163:903–21
128. Wabuke-Bunoti, M. A. N., Fan, D. P. 1983. *J. Immunol.* 130:2386–91
129. Wabuke-Bunoti, M. A. N., Taku, A., Fan, D. P., Kent, S., Webster, R. G. 1984. *J. Immunol.* 133:2194–201
130. Yamada, A., Young, J. F., Ennis, S. A. 1985. *J. Exp. Med.* 162:1720–25
131. Braciale, T. J., Henkel, T. J., Lukacher, A., Bracialc, V. L. 1986. *J. Immunol.* 137:995–1002
132. Scholtissek, C., Rohde, W., von Hoyningen, V., Rott, R. 1978. *Virology* 87:13–20
133. Laver, W. G., Webster, R. G. 1973. *Virology* 51:383–91
134. Ward, C. W., Dopheide, T. A. 1981. *Biochem. J.* 195:337–40
135. Fang, R., Min Jou, W., Huylebrook, D., Devos, R., Fiers, W. 1981. *Cell* 25:315–23

136. Daniels, R. S., Skehel, J. J., Wiley, D. C. 1985. *J. Gen. Virol.* 66:457–64
137. Gething, M. J., Bye, J., Skehel, J. J., Waterfield, M. D. 1980. *Nature* 287:301–6
138. Verhoeyen, M., Fang, R., Min Jou, W., Devos, R., Huylebrook, D., et al. 1980. *Nature* 286:771–76
139. Yewdell, J. W., Webster, R. G., Gerhard, W. U. 1979. *Nature* 279:246–48
140. Natali, A., Oxford, J. S., Schild, G. C. 1981. *J. Hyg.* 87:185
141. Wang, M. L., Skehel, J. J., Wiley, D. C. 1986. *J. Virol.* 57:124–28