1986

# Characterization Of The Active Site Of The Influenza Virus Neuraminidase Protein (Mutagenesis).

Michael Ronald Lentz
*University of Alabama at Birmingham*

**Lentz, Michael Ronald**

CHARACTERIZATION OF THE ACTIVE SITE OF THE INFLUENZA VIRUS NEURAMINIDASE PROTEIN

*The University of Alabama in Birmingham* PH.D. 1986

**PLEASE NOTE:**

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs or pages ✓
2. Colored illustrations, paper or print _____
3. Photographs with dark background ✓
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Dissertation contains pages with print at a slant, filmed as received _____
16. Other _____

University
Microfilms
International

# CHARACTERIZATION OF THE ACTIVE SITE OF THE INFLUENZA VIRUS NEURAMINIDASE PROTEIN

by

MICHAEL RONALD LENTZ

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Microbiology
in The Graduate School, The University of Alabama at Birmingham

BIRMINGHAM, ALABAMA

1986

ABSTRACT OF DISSERTATION
GRADUATE SCHOOL, UNIVERSITY OF ALABAMA AT BIRMINGHAM

Degree Doctor of Philosophy Major Subject Molecular Cell Biology

Name of Candidate Michael R. Lentz

Title Characterization of the Active Site of the Influenza Virus Neuraminidase Protein

There is presently no effective means of controlling the spread of influenza virus. This is largely the result of antigenic variation in the surface glycoproteins of the virion, the hemagglutinin (HA) and the neuraminidase (NA). This variation allows new strains of the virus to arise in the population and escape the established immunity of the host to previous strains. Antigenic variation also reduces the effectiveness of vaccines against new strains of the virus. A more effective means of controlling this virus may be to target drugs or inhibitors to more conserved regions of viral proteins.

The influenza virus NA protein possesses enzyme activity which is highly conserved in all strains despite 50-60% variation in amino acid sequences of this protein. The enzyme facilitates movement of the virus to and from the site of infection in the host by cleaving terminal sialic acid residues from glycoproteins. The conserved nature of this enzyme activity makes the active site of the NA a possible target for antiviral therapy. This dissertation describes experiments designed to further our understanding of the active site of this enzyme.

A full-length cDNA clone of the NA gene from the A/Tokyo/3/67 strain of influenza was synthesized from viral RNA and sequenced. The cDNA was then inserted into an SV40 virus recombinant expression vector designed and constructed to facilitate expression of the cloned NA gene. The vector expressed NA which was indistinguishable from NA from influenza virions; the protein was transported efficiently to the infected cell surface and was enzymatically active.

The NA gene was subsequently cloned into an M13 phage vector and used as template for site-specific oligonucleotide-directed mutagenesis. Amino acids which are conserved and located in the substrate binding pocket were chosen as targets for a total of twelve mutations, ranging from very conservative substitutions to more dramatic changes. The mutated NA genes were cloned back into the SV40 expression vector, and synthesis, transport, and enzyme activity of the NA proteins in infected cells was assayed.

All but two of the mutant proteins were indistinguishable from wild-type in their expression, transport and folding. The enzyme activity of these mutants ranged from fully active in the case of one active site mutant and one control, to completely inactive in others. Several mutants had intermediate activity and the effect of pH on these enzymes was determined. The results obtained suggest a possible mechanism for neuraminidase catalytic activity which has features in common with the mechanism of lysozyme. It is likely that histidine 274 elevates the $pK_a$ of nearby glutamic acid 276 which is then able to donate a proton to the glycosidic bond. Other conserved residues in the active site pocket are likely to be involved in binding the substrate or stabilizing a charged intermediate.

This information may now be used in a more rational approach to design drugs to inhibit the NA activity of the influenza virus. Such drugs may control the spread of the virus regardless of its antigenic properties.

Abstract Approved by: Committee Chairman _____

Program Director _____

Date 8/22/86 Dean of Graduate School _____



PS-5744

ACKNOWLEDGMENTS

A large group of people take part in a project of this type and I am grateful to everyone who has helped make this project possible; however, a number of individuals deserve special recognition.

First and foremost I would like to thank my advisor, Gillian Air, for her guidance, patience, and enthusiasm over the last four years. Gillian provided the laboratory space, funding, and supplies that made this project possible, and taught me the laboratory, writing, and speaking skills I hope to use for the rest of my career. I am especially grateful for the opportunities she gave me to travel and meet other scientists and present my own work at a variety of meetings.

My thesis committee was excellent and provided valuable assistance throughout this project. I had many helpful discussions with Eric Hunter who also provided vector DNA and taught me the basics of tissue culture. Dick Compans made numerous helpful suggestions and provided training grant support that also allowed me to travel. Two of my committee members made the special effort necessary to be at my committee meetings and defense from off the UAB campus, Rob Webster from Memphis and Bud Suddath from Atlanta. Rob also provided all of the antibody preparations and made many helpful suggestions. Bud was extremely helpful in making sense of the protein data and in deciding what mutations would be the most interesting.

I consider myself very lucky to have had the opportunity to meet and talk on a number of ocassions with Drs. Greame Laver and Peter Colman from Australia. Both were instrumental in analyzing the structure of neuraminidase which made this whole project possible. Greame also provided influenza virus stocks and Peter provided figures, slides, and critical help in analyzing data.

Closer to home, I would like to thank Dr. Jeff Engler for synthesizing most of the oligonucleotides used for the mutagenesis and Dr. John Wills, formerly of Eric Hunter's

lab, for help in getting started in tissue culture. I am also grateful for the other members of Eric's lab for help and friendship and for letting me borrow things when I couldn't find them in our lab.

In Gillian's lab, Terri Shipman, Rodney Tucker, Dan Lakeman, Lynn Ritchie, and Cindy Miller have provided expert assistance. The other graduate students in Gillian's lab, Maryna Els and Pat Bossart, helped me realize how valuable a friend can be.

Finally, I would like to thank my wife, Dorothy for her patience, love, and friendship; and my parents, whose love, inspiration, and support have helped me more than they could ever know.

TABLE OF CONTENTS

Table of Contents (Continued)

Table of Contents (Continued)

LIST OF TABLES

LIST OF FIGURES

SITE-DIRECTED MUTATION OF THE ACTIVE SITE OF INFLUENZA NEURAMINIDASE AND IMPLICATIONS FOR THE CATALYTIC MECHANISM



SUMMARY AND DISCUSSION

## LIST OF ABBREVIATIONS

| | |
|---|---|
| A | the adenine base of a nucleotide |
| Å | angstrom (s) |
| Ab | antibody molecule (s) |
| A549 | absorbance at wavelength 549 nanometers |
| AMV | avian myoblastosis virus |
| Arg | amino acid arginine |
| Asn | amino acid asparagine |
| Asp | amino acid aspartic acid |
| C | the cytidine base of a nucleotide |
| C-terminus | the carboxy terminal region of a polypeptide |
| cDNA | complementary deoxyribonucleic acid |
| CMS | calcium, magnesium, saline buffer |
| CPE | cytopathic effect (s) |
| CV-1 | African green monkey kidney cell line |
| DEAE | diethylaminoethyl |
| DNA | deoxyribonucleic acid |
| dATP | deoxyadenosine triphosphate |
| dCTP | deoxycytidine triphosphate |
| dGTP | deoxyguanosine triphosphate |
| dTTP | deoxythymidine triphosphate |
| EDTA | disodium ethylenediamine-tetraacetic acid |
| ELISA | enzyme-linked immunosorbent assay |
| f. p. | fusion peptide |

List of Abbreviations (continued)

| | |
|---|---|
| G | guanosine base of a nucleotide |
| G70C | influenza virus A/tern/Australia/G70C/75 (H11N9) |
| Glu | amino acid glutamic acid |
| Gly | amino acid glycine |
| HxNy | influenza virus of hemagglutinin subtype x and neuraminidase subtype y |
| HA | hemagglutinin glycoprotein of influenza virus |
| HAU | hemagglutinating units |
| His | amino acid histidine |
| hr | hour (s) |
| Ile | amino acid isoleucine |
| Km | Michaelis constant |
| Leu | amino acid leucine |
| Lys | amino acid lysine |
| M | molar |
| mA | milliampere (s) |
| min | minute (s) |
| mM | millimolar |
| mm | millimeter (s) |
| mol wt | molecular weight |
| M1 | the membrane or matrix protein of influenza virus |
| M2 | nonstructural glycoprotein of influenza virus |
| mRNA | messenger ribonucleic acid |
| N-linked | asparagine-linked |
| N-terminus | the amino terminal region of a polypeptide |
| NA | neuraminidase glycoprotein of influenza virus |
| NAG | N-acetylglucosamine |

List of Abbreviations (continued)

| | |
|---|---|
| NANA | N-acetylneuraminic acid |
| NANL | N-acetylneuraminyl lactose |
| nm | nanometer (s) |
| nmole (s) | nanomole (s) |
| NP | nucleoprotein of influenza virus |
| NS1 | large nonstructural protein of influenza virus |
| NS2 | small nonstructural protein of influenza virus |
| PA | polymerase protein of influenza virus |
| PB1 | polymerase protein of influenza virus |
| PB2 | polymerase protein of influenza virus |
| PBS | phosphate-buffered saline |
| Phe | amino acid phenylalanine |
| RER | rough endoplasmic reticulum |
| REV | revertant |
| RF | replicative form DNA |
| RNA | ribonucleic acid |
| RNase | ribonuclease |
| rpm | revolutions per minute |
| SDS | sodium dodecyl sulfate |
| Ser | amino acid serine |
| SV40 | simian virus 40 |
| SVNAwt | recombinant SV40 vector expressing wild-type neuraminidase |
| SVNA178L | recombinant SV40 vector expressing neuraminidase with a substitution of leucine for tryptophan at residue 178 |
| T | thymidine base of a nucleotide |
| TBE | tris/borate/EDTA buffer |
| Thr | amino acid threonine |

List of Abbreviations (continued)

| | |
|---|---|
| tRNA | transfer ribonucleic acid |
| Tokyo/67 | influenza virus strain A/Tokyo/3/67 (H2N2) |
| Trp | amino acid tryptophan |
| Tyr | amino acid tyrosine |
| U | unit (s) |
| U/I | uninfected |
| Val | amino acid valine |
| VH | virus stock of recombinant SV40 vector with no NA gene plus helper virus |
| Vmax | maximum velocity |
| vRNA | viral ribonucleic acid |
| v/v | volume to volume ratio |
| WSN | influenza virus strain A/WSN/33 (H1N1) |
| WT | wild-type |
| µg | microgram (s) |
| µl | microliter (s) |

# INTRODUCTION

Influenza continues to have a major impact on the health and economy of the world despite many important recent advances in vaccine and drug therapy that have virtually eliminated some virally induced diseases and reduced the incidence of many more. Current influenza vaccines rapidly lose their effectiveness as new strains of the virus appear in the population. In an average season, influenza accounts for approximately 10,000 to 20,000 excess deaths, primarily among the elderly and the very young. In a particularly severe outbreak of the disease, this figure can be as high as 40,000 to 70,000 deaths. The pandemic of 1918-1919 killed 20-30 million people worldwide. Estimated direct and indirect costs to the economy of a severe influenza epidemic can be as high as $3 billion (Murphy and Webster, 1985), due largely to morbidity in the working population.

The Influenza Virus Structure, RNA, and Proteins. The structure of the virus has been studied extensively in an effort to understand how the virus is able to elude current vaccine strategies. Virus particles are highly pleomorphic, ranging from spherical particles with a diameter of approximately 120 nm to highly elongated forms usually isolated from primary infections. Prominant glycoprotein spikes are seen projecting from the surface of the viral envelope. The envelope is lined by a matrix protein shell which surrounds the helical nucleocapsids of the genome. The structure of the virus is illustrated in Figure 1.

Influenza is a member of the Orthomyxoviridae family which is characterized by a segmented RNA genome of negative polarity (genome RNA complementary to mRNA). In influenza A viruses, there are eight genome segments, each coding for one or more proteins. These segments range from 890 to 2431 nucleotides; all share common sequences of 13 nucleotides at the 5' end and 12 nucleotides at the 3' end (Skehel and Hay, 1978; Robertson, 1979; Desselberger et al., 1980). The conserved sequences show partial complementarity and may be important for replication of the RNA segments (Robertson, 1979; Desselberger et al., 1980).

The three largest RNA segments code for the three polymerase proteins, PA, PB1, and PB2. PB1 and PB2 of strain A/PR/8/34 are basic proteins with a net charge of +28 at pH 6.5 while PA is acidic with a -13.5 charge at pH 6.5 (Winter and Fields, 1982). PB2 functions by recognizing the cap structure on cellular mRNAs (Ulmanen et al., 1981). These caps are cleaved from the cellular mRNA and serve as a primer for synthesis of viral mRNA. It is not clear at this time if PB2 is also responsible for this cleavage reaction. Crosslinking experiments performed during the initial steps of viral mRNA synthesis have established PB1 as the protein responsible for initiating elongation of the viral message (Ulmanen et al., 1981). PB1 may also be associated at least in part in continued elongation of the mRNA. Studies of temperature-sensitive mutants defective in segment three indicate that the PA protein is involved in the synthesis of new viral RNA in infected cells (Krug et al., 1975). It does not appear to play a role in virus-specific mRNA synthesis.

RNA segment four encodes the hemagglutinin glycoprotein that is located on the surface of the virus. The structure and function of this protein will be discussed in detail later.

The nucleocapsid protein (NP) is coded for by the fifth gene segment. This protein associates closely with the viral RNA and is the major protein component of the ribonucleoprotein complexes (Pons et al., 1969; Compans et al., 1972). NP is also closely associated with the PA, PB1, and PB2 proteins to form the transcription complex although there is no clear evidence from temperature-sensitive mutants that NP is directly involved in transcription. In fact, transcriptionally active complexes have been isolated which do not contain NP protein (Kawakami and Ishihama, 1983). NP also appears to interact with the matrix protein (Rees and Dimmock, 1981) and may play an important role in the maturation of the viral particles by directing the ribonucleoprotein complexes in the infected cell to the site of budding at the plasma membrane. NP is one of the type-specific antigens used to classify type A, B, and C influenza viruses.

Gene segment six codes for the neuraminidase surface glycoprotein of the virus. The structure and function of neuraminidase will be discussed in detail below.

Two proteins are coded for by segment seven, the matrix or membrane protein (M1) and the non-structural protein, M2. M1 is the most abundant protein of the virus and interacts with the lipid envelope to form a protein shell on the internal surface of the membrane (Compans et al., 1970, 1972; Schulze, 1972). M1 may provide a recognition site on the inner surface of the plasma membrane for the ribonucleoprotein complexes during virus maturation (Choppin et al., 1972). M1 is the second type-specific antigen used to distinguish the influenza A, B, and C viruses (Schild, 1972; Oxford and Schild, 1976).

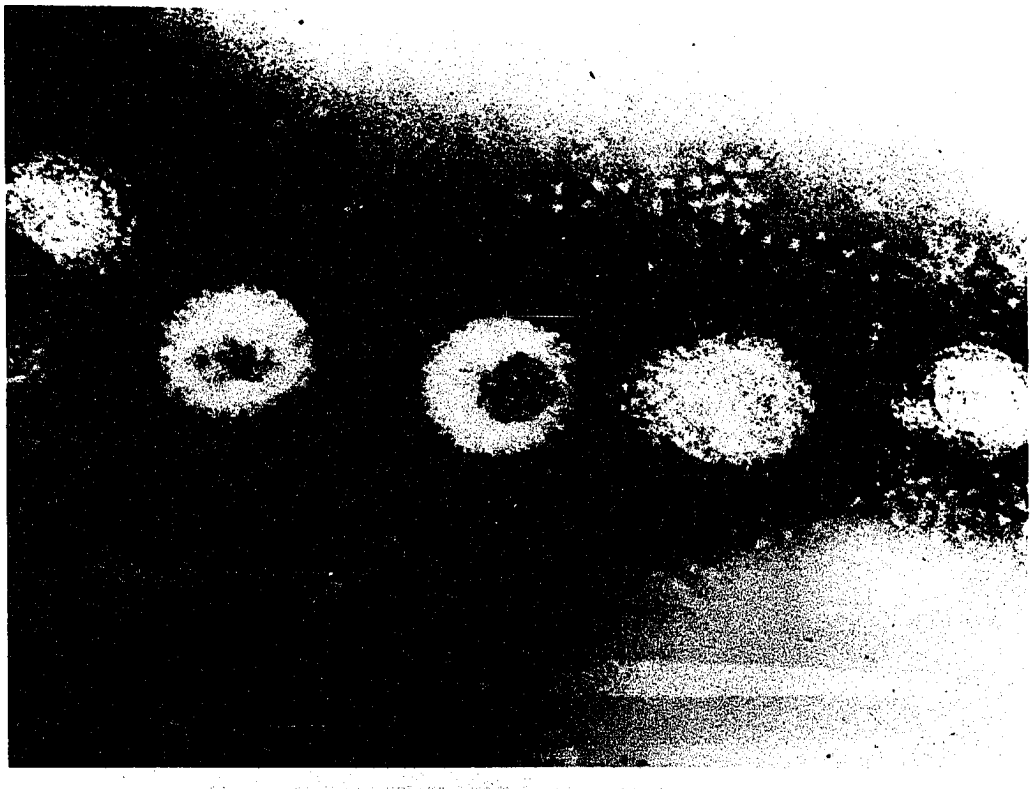
M2 is a recently discovered integral membrane protein found on the surface of infected cells but probably not in influenza virions. The protein is synthesized from a spliced message from genome segment seven (Lamb and Choppin, 1981; Palese et al., 1981). M2 on the infected cell surface is recognized by antibodies and may be a major target antigen for cross-reactive cytotoxic T cells (Lamb et al., 1985).

Two nonstructural proteins, NS1 and NS2, are synthesized from RNA segment eight. NS1 is probably involved in the shutoff of virion RNA synthesis as well as host cell protein synthesis (Lazarowitz et al., 1971; Compans, 1973; Wolstenholme et al., 1980; Koennecke et al., 1981). NS1 is a phosphoprotein, with one or two threonines per molecule modified by a phosphate (Privalsky and Penhoet, 1977, 1978, 1981; Almond and Felsenreich, 1982).

NS2 is synthesized from a small spliced mRNA from segment eight (Lamb and Choppin, 1979; Inglis et al., 1979). NS2 is found in the cytoplasm of infected cells late in infection but its function is not known (Lamb et al., 1978; Lamb and Lai, 1980; Mahy et al., 1980). The gene segments and proteins of influenza are summarized in Table 1.

Antigenic Variation. The hemagglutinin (HA) and neuraminidase (NA), coded for by segments four and six respectively, are the surface glycoproteins of the virus. Both of these proteins are anchored to and project from the viral envelope and are the major

Figure 1. Structure of the influenza virus. (Top) Electron micrograph of virus particles. Glycoprotein spikes are clearly visible on the surface of the virion. (Bottom) Schematic diagram of the virus particle. Glycoproteins are embedded in the viral envelope, which is lined by a matrix shell. Within the virion are eight ribonucleoprotein complexes, each containing a RNA genome segment coding for one or two proteins.



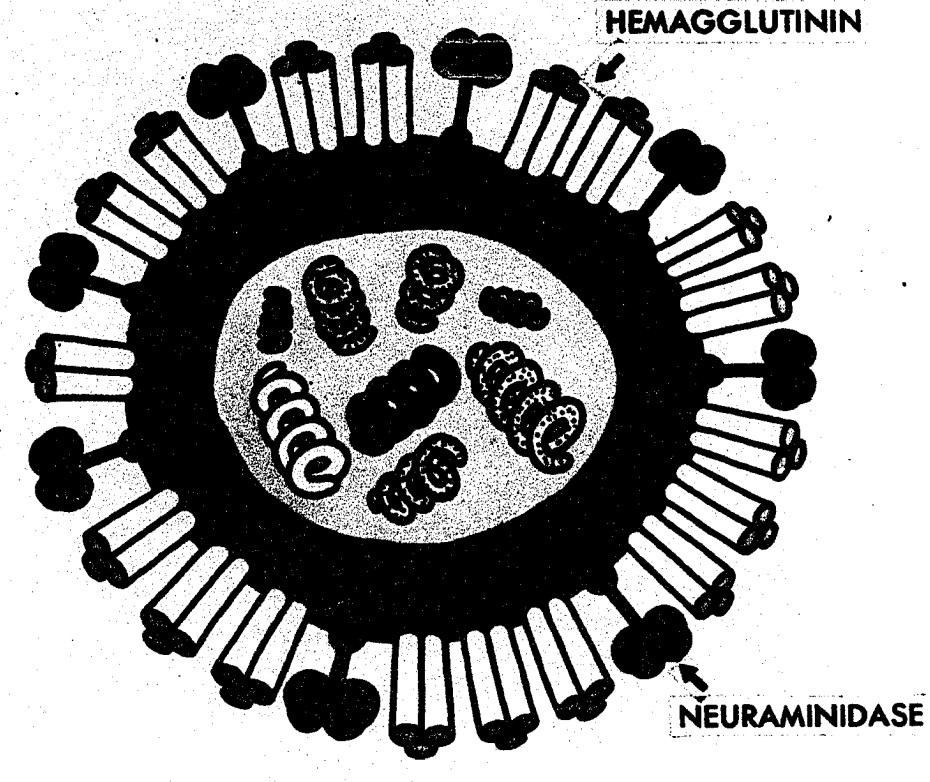


TABLE 1
INFLUENZA VIRUS GENOME RNA SEGMENTS
AND CODING ASSIGNMENTS[a]

| Segment | Length (nucleotides) | Encoded polypeptide | Nascent polypeptide length (aa) | Mature polypeptides | Approx. No. molecules per virion | Remarks |
|---|---|---|---|---|---|---|
| 1 | 2,341 | PB2 | 759 | | 30-60 | Host cell RNA cap binding: component of RNA transcriptase. |
| 2 | 2,341 | PB1 | 757 | | 30-60 | Initiation of transcription: component of RNA transcriptase. |
| 3 | 2,233 | PA | 716 | | 30-60 | Component of RNA transcriptase. |
| 4 | 1,778 | HA | 566 | HA1 326<br>HA2 222 | 500 | Surface glycoprotein, receptor-binding and fusion activity: Trimer: Major antigenic determinant. |
| 5 | 1,565 | NP | 498 | | 1000 | Associated with RNA segments to form ribonucleoprotein: Structural component of RNA transcriptase. Type-specific antigen. |
| 6 | 1,413 | NA | 454 | 454 | 100 | Surface glycoprotein; neuraminidase activity: Tetramer: Surface antigen. |
| 7 | 1,027 | M | 252 | 252 | 3000 | Major protein component of virus: underlies lipid bilayer: Type-specific antigen. |
| | | M2 | 96 | | | Spliced mRNA, non-structural protein: function unknown. |
| | | ? | ?9 | | | Spliced mRNA, peptide predicted by nucleotide sequence only. |
| 8 | 890 | NS1 | 230 | | | Non-structural protein: shut-off of host cell protein and viral RNA synthesis. |
| | | NS2 | 121 | | | Spliced mRNA, non-structural protein: function unknown. |
| | 13,588 | | | | | |

[a] For A/PR/8/34 strain: See Lamb (1983) and Krug (1983) for full references.

antigens of the virus. The HA and NA surface glycoproteins undergo antigenic variation which is largely responsible for the ability of the virus to escape the host's immune system as well as available vaccines. Two distinct types of antigenic variation have been observed for influenza A viruses, antigenic drift and antigenic shift. Antigenic drift consists of relatively minor changes in the antigenic structure of the surface proteins that accumulate over time. These changes result from random point mutations that occur in the genes for the HA and NA. Some of these mutations may interfere with the biological function of the protein and are therefore lethal and never seen, while others may change the antigenic structure of the protein and give the resulting variant a selective advantage in overcoming the immunological resistence of the host. These variants are responsible for the moderate seasonal outbreaks of influenza.

Antigenic shift has only been observed for influenza A viruses and results in much more dramatic antigenic changes in the virus. Antigenic shift occurs when a new subtype of influenza A appears in the human population replacing the type that had previously been circulating. The new virus has an antigenically unrelated HA and sometimes NA from the subtype that had been circulating prior to the shift. Antigenic shift results in the major pandemics of influenza. Since the influenza virus was first isolated in 1933, three major shifts have occurred. In 1957, the H2N2 subtype (Asian influenza) replaced the H1N1 subtype that had been circulating since 1933. In 1968, only the HA shifted when the H3N2 (Hong Kong influenza) replaced the Asian flu. The H1N1 subtype (Russian influenza) reappeared in 1977.

Three mechanisms have been postulated for the emergence of new subtypes in antigenic shift. First, there is evidence that some strains of the virus disappear after circulating in the human population and then reappear many years later. The Russian flu (H1N1) that reappeared in China in 1977 seems to be identical in all respects to a virus that was causing epidemics in 1950 (Nakajima et al., 1978; Scholtissek et al., 1978). Where this virus was for 27 years remains a mystery.

Second, the segmented genome of the influenza virus allows recombination (reassortment) events to readily occur when one cell becomes infected with two different influenza viruses. Thus, the antigens in a new strain might be derived from an avian or other mammalian strain. Such viruses cannot normally productively infect humans, but through a genetic reassortment event a human strain with a new antigenic structure could result. The HA from the Hong Kong (H3N2) strain that emerged in 1968 is closely related to that of avian (A/Duck/Ukraine/63) as well as equine (equi-2) viruses while the rest of the genes are apparently derived from the human Asian (H2N2) strain (Laver and Webster, 1973; Fang et al., 1981; Ward and Dopheide, 1981).

Finally, it may be possible for an influenza virus from a non-human animal host to directly mutate to a virus with the capacity to infect humans. The host range determinants for different animal viruses are not understood, but recent evidence suggests that the ability of an influenza virus to infect different hosts can vary. A virus closely related to fowl plague virus [A/FPV/Dutch/27 (H7N7)] previously not found in mammals, killed 20% of the harbour seal population of the northeast coast of the United States in 1979-1980. The seals had a severe respiratory infection and high concentrations of the influenza virus were found in the brains and lungs of the seals (Webster et al., 1981). Such an event could conceivably occur in humans instead of seals.

The structures of the HA and NA have been studied in detail in an effort to understand the mechanisms of antigenic shift and drift. Both proteins have been crystallized, their three-dimensional structures determined and the nucleotide and amino acid sequences for a large number of strains for both proteins are now known.

Structure and Function of the Hemagglutinin. The HA protein accounts for about 25% of the protein of the virus and is the major antigen to which neutralizing antibodies (Ab) are directed (Laver and Kilbourne, 1966). The HA has two major functions during the infectious cycle of the virus. It is responsible for the initial attachment of the virus to the host cell by binding to specific, sialic acid containing receptors on the cell surface

(Hirst, 1942), and is responsible for initiating the infection (Klenk et al., 1975; Lazarowitz and Choppin, 1975) by fusing viral and host cell membranes at low pH.

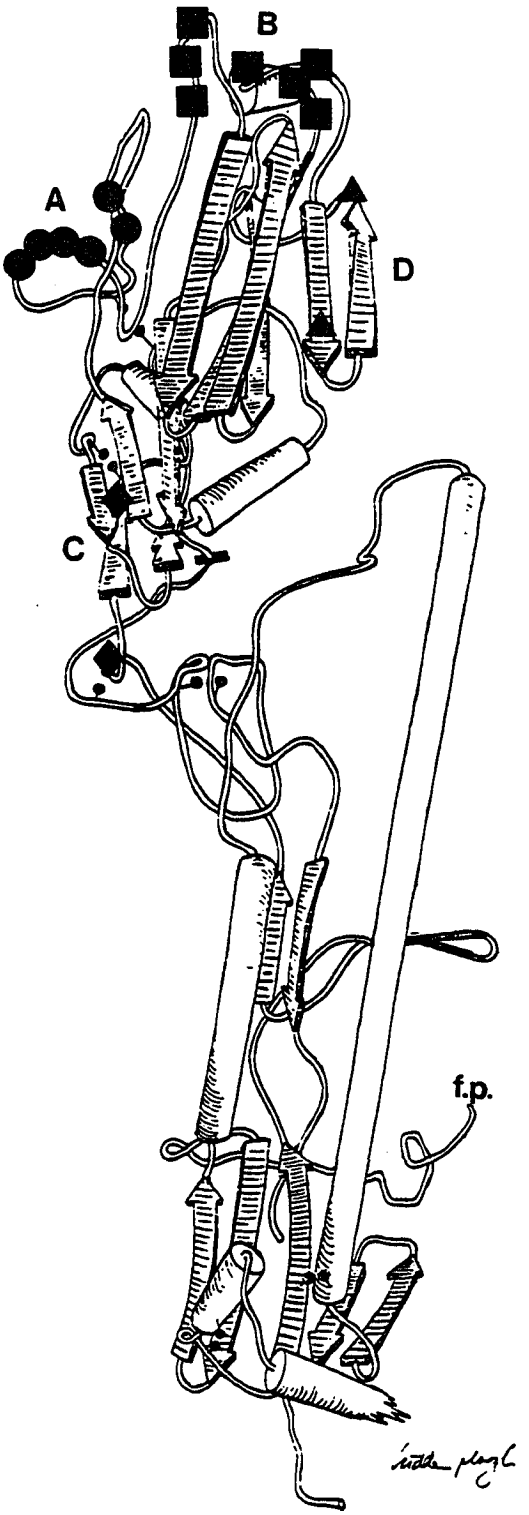
The protein is synthesized on ribosomes bound to the rough endoplasmic reticulum (RER) (Blobel, 1980) and transported through the RER to the Golgi apparatus during which time a number of post-translational modifications are made. An N-terminal translocation signal peptide is removed (Air, 1979; McCauley et al., 1979) and N-linked oligosaccharides are attached (Compans, 1973; Hay, 1974; Klenk et al., 1974; Elder et al., 1979; McCauley et al., 1980). The protein remains attached to the membrane by a hydrophobic anchor sequence near its C-terminus.

An internal cleavage of the HA into two disulfide-linked polypeptides, HA1 and HA2, is required for infectivity of the virus (Lazarowitz et al., 1975; Klenk et al., 1975; Huang et al., 1980). This cleavage event exposes a highly hydrophobic peptide at the N-terminus of HA2 which mediates the fusion of the viral and host cell membranes.

The three-dimensional structure of the bromelain released HA of the Hong Kong strain (H3N2) revealed that the protein is a trimer, each polypepteide consisting of a long α-helix stem (HA2) to which is attached a globular head (HA1) (Wilson et al., 1981; Wiley et al., 1981). The sialic acid receptor binding site is located on the globular region which is distal to the viral envelope, as are four distinct antigenic sites, labeled A, B, C, and D. The fusion peptide is located 100 Å from the distal tip of the HA, so that an acid induced conformational change is required to bring the fusion peptide in proximity to the cell membrane within the lysozome (Daniels et al., 1985). The structure of the HA is diagrammed in Figure 2.

The genes of a number of subtypes and strains of HA have been sequenced revealing information about the distribution of antigenic sites on the surface of the molecule (reviewed by Ward, 1981). Functionally active HA protein has been synthesized from cloned HA genes in recombinant eukaryotic expression systems (Gething and Sambrook, 1981), and expression of mutant HA genes in these and other systems have been used to study the general characteristics of cell surface glycoproteins as well as their pathways of

Figure 2. Structure of the influenza hemagglutinin. The structure of a single monomer is shown. HA1 is the globular region at the top of the molecule; HA2 is the α-helical region at the bottom. Positions of amino acid substitutions in monoclonal variants define antigenic sites A (●), B (■), C (◆), and D (▲). The receptor binding site is located at the distal tip of the molecule, and the location of the fusion peptide (f.p.) exposed by activational cleavage is shown. Drawing by Hidde Pleogh. Data from Wilson et al., 1981; Wiley et al., 1981; Laver et al., 1979; Daniels et al., 1983; and Newton et al., 1983.



cellular synthesis and transport (Gething and Sambrook, 1982). Details of biological functions of the molecule have been studied using naturally occurring and in vitro synthesized mutants of HA (Daniels et al., 1985; Doms et al., 1986; Rogers et al., 1983, 1985). These and many other experiments have made the HA protein of influenza one of the best understood viral proteins.

Structure, Function, and Activity of the Neuraminidase. The neuraminidase (NA) is the other integral membrane surface glycoprotein of the virus and accounts for 5-10% of the virion protein. It too has been studied in detail, although it is not as well understood as the HA. It is synthesized on ribosomes bound to the RER and transported to the Golgi apparatus during which time the protein is glycosylated. No post-translational cleavage events occur and even the initiating methionine is retained (Blok et al., 1982). There is a highly conserved sequence of six polar amino acids at the N-terminus of the protein (Blok and Air, 1982; Air et al., 1985b) followed by a sequence of hydrophobic amino acids which serves both as a signal to initiate cotranslational translocation of the polypeptide into the RER and to anchor the protein in the membrane (Fields et al., 1981). Removal of the hydrophobic N-terminal region of NA expressed from a cloned gene further demonstrates that this region of the molecule is critical for translocation, glycosylation and surface expression (Davis et al., 1983; Markoff et al., 1984). This is in contrast to the HA and the majority of other membrane bound proteins which are anchored by their C-terminus and have a proteolytically cleaved signal peptide at their N-terminus. Immunoelectron microscopic studies of influenza virions have demonstrated that the NA may occur on the virus particle in discrete patches unlike the HA which appears to be evenly distributed on the virus surface (Compans et al., 1969; Murti and Webster, 1986).

NA is an enzyme (acylneuraminyl hydrolase, EC 3.2.1.18) which cleaves the α–ketosidic linkage between a terminal sialic acid (N-acetylneuraminic acid) residue and the adjacent sugar on polysaccharides as shown in Figure 3 (Gottschalk, 1957). Influenza NA preferentially cleaves α(2,3) linkages (Corfield et al., 1982, 1983; Paulson et al., 1982)

and is a true exoenzyme since internal sialic acid in the correct linkage will not be cleaved by the enzyme (Drzeniek, 1972). The best substrates contain an acetyl group at the nitrogen atom, although a glycolyl group at this position will also be cleaved. Influenza NA is also active towards O-acetylated substrates. The negative charge of the carboxyl group appears to be very important, as substitution at this position prevents cleavage by the enzyme (Drzeniek, 1973).

The role of the NA during the course of viral infection is not well understood. Studies over the past 40 years indicate that it probably has a general function of facilitating the movement of the virus within the host which can be accomplished in a number of ways. The virus enters through the nasal route and encounters a mucosal layer protecting the target respiratory epithelial cells. The mucous layer consists largely of highly sialated glycoproteins (Schauer, 1985) and the neuraminidase may aid in its penetration to allow access of the virus to the target cells (Burnet et al., 1947; Burnet, 1948). The NA may also facilitate movement of the virus from the infected cell by cleaving the sialic acid receptor for the HA from the cell surface and from the carbohydrate side chains of HA and NA on the newly synthesized virus particles, preventing their aggregation (Seto and Rott, 1966; Palese et al., 1974; Palese and Compans, 1976; Griffin and Compans, 1979; Basak et al., 1985). Cleavage of sialic acid from HA may facilitate the proteolytic cleavage event required for activation of the HA (Schulman and Palese, 1977).

Recently, a reassortant virus was described which apparently does not incorporate active NA into virus particles, but which is fully infectious for chickens (Breuning and Scholtissek, 1986). This virus has the PB2 and NA genes of A/Hong Kong/1/68 and the remaining genes from A/fowl plague/Rostock/34. At 33°, NA is overproduced, transported to the cell surface, enzymatically active, and incorporated into virions, but at 40°, little NA activity can be detected and the protein is blocked in transport at the RER. Since the virions produced at 40° are fully infectious for chickens (body temperature 41°), these results raise questions about the role of the NA in the life cycle of the virus. The authors suggest that viral NA may be involved only in penetrating the mucous layer during the initial infection.

If this were the case, the NA active site would not be a very suitable target for antiviral therapy. These conclusions are open to question for a number of reasons. First, the virions produced at 40° are never completely without NA or NA activity. NA protein may be present at 3% of that at 33°, while the NA activity in some experiments was over 30% of the activity at 33°. The activity of NA from natural isolates of different subtypes can vary greatly, and this is also true of laboratory recombinants as shown by the authors. Viral NA activity may only be required in low levels to accomplish the functions described above; the reassortant assayed in the paper may have enough activity. Secondly, cellular NA may be more active in chicken cells and make up for the apparent reduction of viral activity. The authors failed to determine whether or not sialic acid was present on the glycoproteins of virions released at 40°. The low viral NA or cellular NA may be enough to remove sialic acid from these proteins, a step believed to be required to prevent progeny virus aggregation. Finally, the authors only analyzed virions produced from primary chick embryo cells in tissue culture. The virions produced in infected chickens may have different NA activity since adult differentiated chicken cells are likely to have very different characteristics from the primary cultures. The effect of antibodies to the Hong Kong NA in both tissue culture and adult chickens infected with the reassortant also needs to be determined. It would seem that the reassortant requires more careful analysis to support or refute the conclusions of the authors. All other data support an important role for the NA during the viral life cycle and it seems likely that inhibition of NA activity may aid in the control of the disease.

The biologically active part of the molecule is a square, box-like "head" which is held above the viral envelope by a long narrow stalk. The heads can be cleaved from the virus by digestion with Pronase and retain all the biological activity of the intact protein. In some strains these heads can be crystallized and the three-dimensional structure of A/Tokyo/3/67 (N2) NA has been determined (Varghese et al., 1983; Colman et al., 1983). The active protein consists of four identical subunits arranged by circular four-fold symmetry with each subunit folded into six, four-stranded ß-sheets. The four subunits are

held together by disulfide bonds (Lazdins et al., 1972), and divalent metal ions bound at the symmetry axis help stabilize the tetramer. This latter activity may be important since ethylenediamine-tetraacetate (EDTA) inhibits enzyme activity and no calcium ions have been seen in the active site (Colman and Ward, 1985). The catalytic site is a deep pocket on the surface of each subunit where sialic acid was shown to bind when soaked into NA crystals (Colman et al., 1983). Four of the five potential carbohydrate attachment sites in the head region are glycosylated (Ward et al., 1982, 1983)

The antigenic structure of the NA has been studied in detail using both naturally occurring field variants and laboratory selected monoclonal variants. Variants were produced by growing the virus in the presence of monoclonal Ab to the NA. Although these Ab do not neutralize the infectivity of the virus, they prevented the release of progeny viruses (Webster et al., 1982, 1984) and allowed variants which no longer bound the Ab to grow through. Variants have been grown and sequenced from A/Tokyo/3/67 (N2) (Laver et al., 1982; Lentz et al., 1984) and compared to variants of X-7F1 (N2) (Air et al., 1985a), and A/tern /Australia/G70C/75 (N9), (Tulloch et al., 1986). Most variants were found to change at a single amino acid which must be responsible for the inability of the Ab to bind to the NA. These sequence changes are located around the rim that circles the substrate binding pocket, with many of the sequence changes concentrated in the outermost side of the protein in the tetramer where access to Ab is greatest. These results are not surprising, since the Ab must almost certainly inhibit enzyme activity in order to prevent the growth of the progeny viruses. Ab which bind to sites on other regions of the molecule will not likely produce monoclonal variants.

The information available on the catalytic site of NA is based on structure analysis and sequence comparison between a variety of NA strains and subtypes. The tertiary structure of the protein, as mentioned above, has a deep depression on the surface of each subunit. Sialic acid, the product of the reaction, was shown to bind in this pocket. A pocket would be expected on an enzyme catalyzing the removal of an end group, rather than the cleft usually found to bind substrate when internal linkages are cleaved (Colman

and Ward, 1985). Analysis of the available NA sequences revealed that 60 amino acids (excluding cysteine) scattered throughout the linear polypeptide are very highly conserved despite up to 50-60% total sequence variation between these strains (Air et al., 1985b; Dale et al., in press). The amino acid sequence alignment illustrating this conservation is shown in Figure 4. In the properly folded protein molecule, 24 of these conserved residues line the pocket on the surface, with their side chains directed into the pocket (Colman et al., 1983; Colman and Ward, 1985). Conservation would be expected for catalytically active amino acids and these conserved residues have therefore been implicated in the enzyme activity of the protein. The conserved residues are Arg 118, Glu 119, Tyr (or Phe) 121, Leu 134, Asp 151, Arg 152, Trp 178, Ser 179, Asp (or Asn) 198, Ile 222, Arg 224, Glu 227, Asp 243, His 274, Glu 276, Glu 277, Arg 292, Asp 293, Asp 330, Lys 350, Thr 365, Arg 371, Tyr 406, and Glu 425 (Colman, 1986). Structural features of the NA are illustrated in Figure 5.

The kinetic behavior of the wild-type Tokyo/67 NA has been studied on the whole virus and on isolated NA heads (Mountford et al., 1982). The pH optimum for the enzyme was found to be 5.8-6.6 for the virus and isolated protein; the kinetics were studied at pH 6.5 using N–acetylneuraminyl lactose as substrate. Sigmoid kinetics were observed for the enzyme over a range of substrate concentrations of 0.03-0.5 mM. The same apparent Km of 0.4 mM was obtained for both high and low substrate concentrations. The $V_{max}$ was found to be 71.4 nmoles sialic acid produced/min/μg of NA for the heads at low substrate concentrations and 36.0 for high concentrations. For the enzyme on the whole virus, the values for $V_{max}$ were 24.5 and 18.5 nmoles sialic acid produced/min/400 HAU virus at low and high substrate concentrations respectively. These parameters were found to vary when different substrates were used in the reaction (Drzeniek, 1972). Complexing the NA with monoclonal antibody abolished the multiphasic behavior of the enzyme. The presence of sigmoid kinetics suggests that cooperative binding of substrate occurs between subunits on a given molecule of the enzyme and that subunit interactions are occurring in the

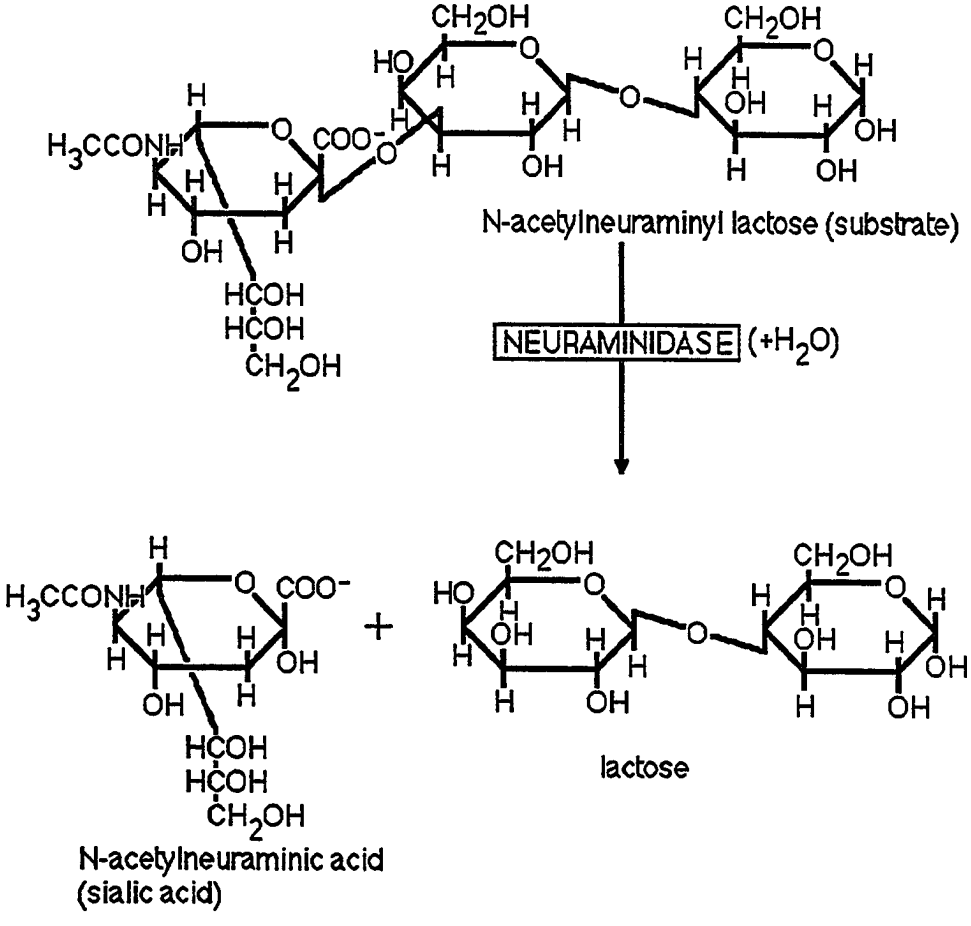



Figure 3. The reaction catalyzed by neuraminidase. N-acetylneuraminyl lactose is used as substrate in this example, however, any polysaccharide to which a terminal sialic acid is linked α(2,3) can serve as substrate. Water is added across the ketosidic bond to release free sialic acid; the remainder of the polysaccharide becomes shortened by one residue.

Figure 4. Amino acid sequence alignment of six subtypes of influenza virus NA highlighting several features. General conserved residues are shaded in green, while conserved residues located in the binding pocket in the folded molecule are shaded blue. Cysteine residues are highlighted in yellow. The amino acids which vary in monoclonal antibody derived antigenic variants of the N2 and N9 subtypes are shaded red. Strains shown are A/Tokyo/3/67 (A/Tok/67), N2; A/Puerto Rico/8/34 (A/PR/8/34), N1; A/equine/Kentucky/1/81 (A/eq/Ken/81), N8; A/equine/Cornell/16/75 (A/eq/Cor/75), N7; A/tern/Australia/G70C/75 (A/G70C/75), N9; and B/Lee/40. Dashes represent spaces inserted to maximize homology. Data from Lentz et al., 1984; Air et al., 1985a; Dale et al., 1986; Colman, 1986; Tulloch et al., 1986; Air et al., unpublished results.

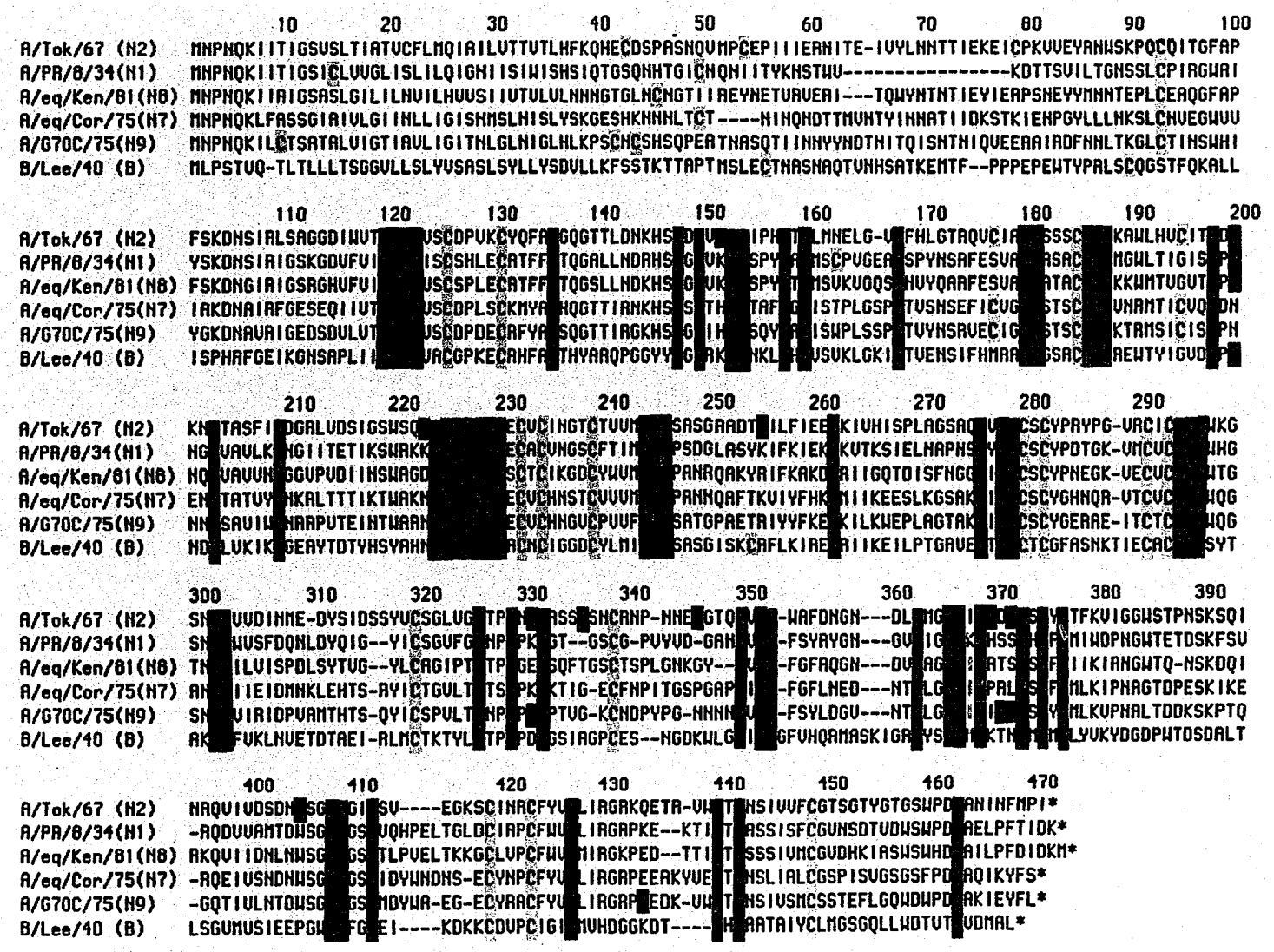
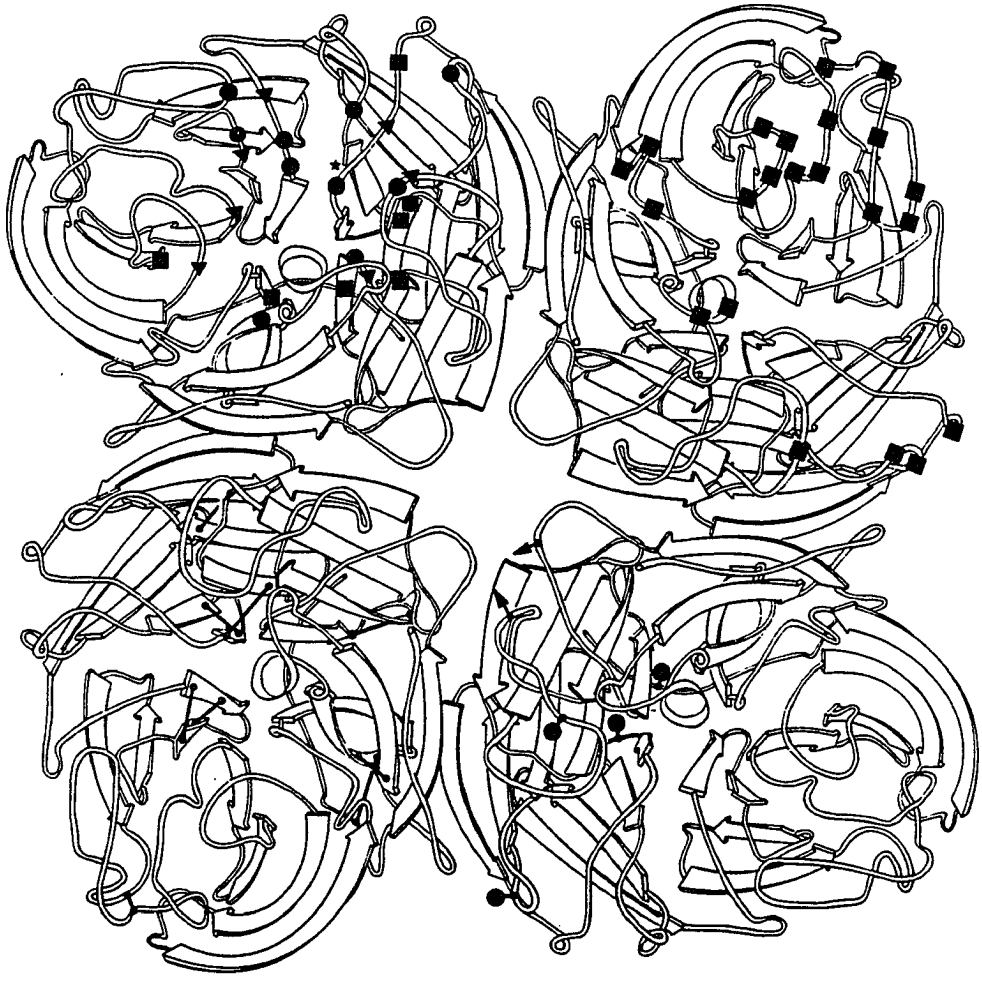


Figure 5. Structure of the influenza N2 neuraminidase. The top view of an tetramer is shown, and each subunit highlights a different feature. Top left, conserved acidic (●), basic (▼) and uncharged (■) residues of the active site. The star marks the location of sialic acid binding. Top right, amino acids found to vary in field isolates of influenza virus. Bottom left, disulfide bonds. Bottom right, glycosylation sites at residues 86, 146, 200, and 234, and metal ligands (arrows). Data from Colman et al., 1983; Varghese et al., 1983; Colman and Ward, 1985; Colman, 1986.



molecule. Complexing the enzyme with antibody prevents this interaction from taking place.

Cellular Neuraminidases and Sialic Acids. Sialic acids and neuraminidases are known to occur in a wide variety of human tissues, particularly brain, liver, mammary tissue, heart, and kidney (Patel, 1978). Sialic acids are often found in high concentrations on the outside of cells, and are components of glycoproteins, gangliosides and polysaccharides. Five main functions have been distinguished for the sialic acids: 1) binding and transport of positively charged compounds and aggregation or disaggregation of cells as a result of their strong electronegative charge ($pK_a$ around 2); 2) influence of the conformation of glycoproteins and their resistance to degradation; 3) antigenic determinants; 4) essential components of receptors; and 5) masking of specific recognition sites (Schauer, 1985). Neuraminidases in various tissues help regulate the sialic acid content of cell and tissue components and influence the contribution of sialic acid to the above functions. Two types of neuraminidases have been observed, a soluble form and a lysosomal form. The pH optimum of the lysosomal form is 4.4 and this enzyme probably removes sialic acid from glycoproteins facilitating their degradation (Patel, 1978). The precise role of the soluble form is not known; it probably has a variety of functions depending on the tissue or cell type in which it is found. Several human disorders are associated with a neuraminidase deficiency, including sialidosis and galactosialidosis (Mueller et al., 1986; Cantz, 1982). These inherited diseases usually result in the accumulation of glycoprotein-derived oligosaccharides. A wide variety of clinical manifestations can result, including opthalmological disorders, skeletal deformities, mental retardation, renal insufficiency, and, in severe cases, death (Cantz, 1982).

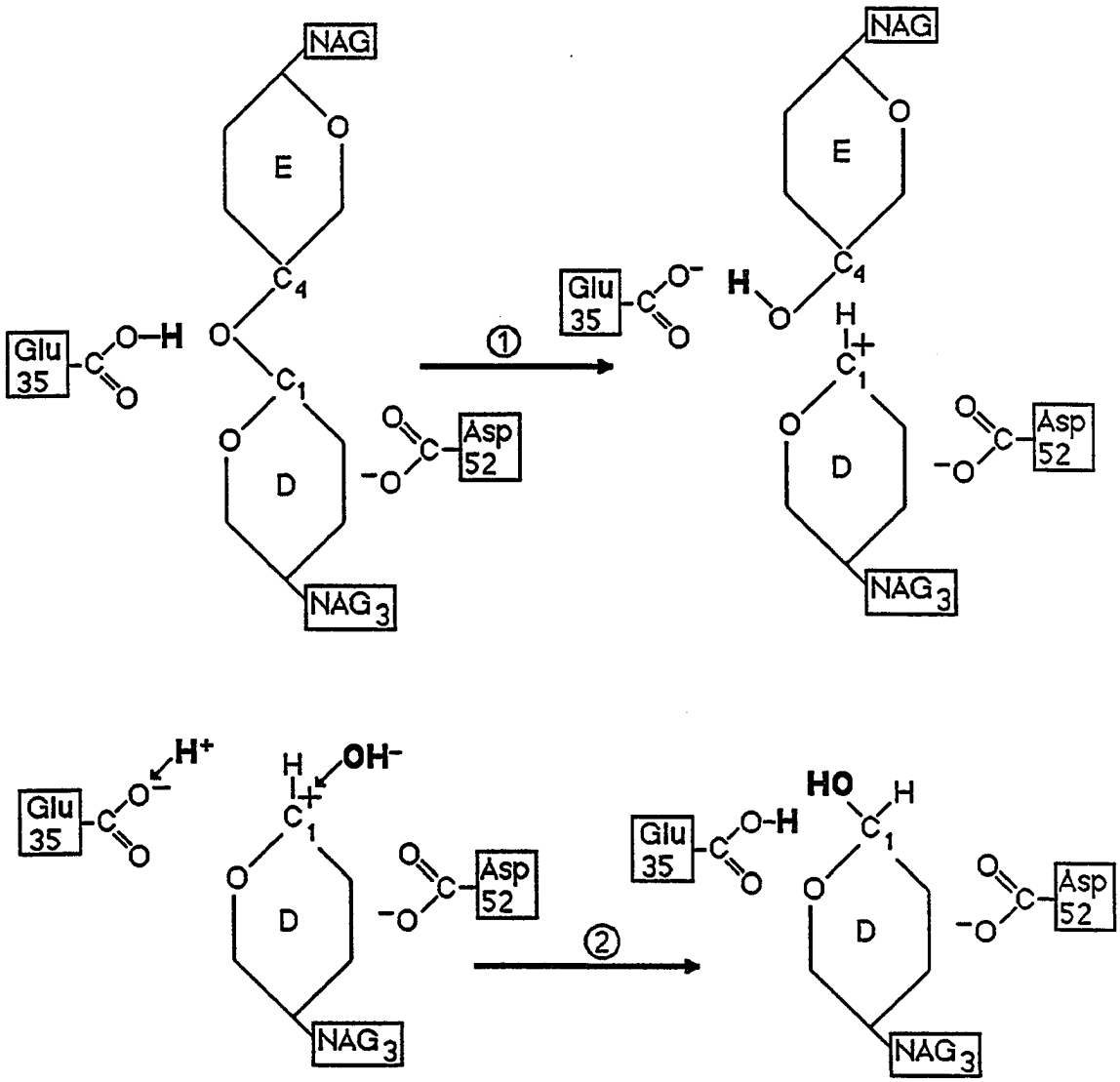
Mechanism of Lysozyme. The catalytic mechanism of a variety of enzymes has been studied in detail, including lysozyme, which also cleaves glycosidic linkages in carbohydrate chains. A mechanism for catalysis was proposed for this enzyme based on the crystal structure which was determined before most of the solution studies of the

enzyme activity were performed. The enzyme binds six sugar residues at a time in six subsites, A through F, and cleavage occurs between the residues occupying sites D and E. Cleavage is general-acid catalyzed by the carboxyl of glutamic acid 35, which remains un-ionized at the pH optimum due to the local hydrophobic environment surrounding the side chain, raising the $pK_a$ to 5.9. This glutamic acid donates its proton to the glycosidic bond. A carboxonium intermediate is formed which is stabilized by ionized aspartic acid 52. A water molecule can then reprotonate Glu 35 and provide a hydroxyl ion to react with the carboxonium ion intermediate and complete the reaction (Phillips, 1967; Blake et al., 1967). This mechanism is shown schematically in Figure 6.

Experimental Approach. The enzyme activity and function of influenza NA is conserved among all the different strains and subtypes and as such the NA may be a suitable target for control of viral infection regardless of the antigenic nature of a particular strain. This dissertation describes research designed to characterize the enzyme active site of this protein. By understanding the mechanism of action of this enzyme and how it interacts with the substrate, it is hoped that a model for the geometry of the active site can be constructed. This model could then be used as a starting point for a rational approach to the design of specific inhibitors able to interfere with the activity of the NA during influenza infection.

This problem has been approached by cloning the gene for the NA of the influenza virus and studying the protein synthesized from the cloned gene. In this way, the NA can be studied without interference from other influenza proteins. By using a cloned cDNA copy of the gene, manipulation of the gene and protein are made much easier than would be the case for a single stranded RNA gene as found in the influenza virion. An SV40 system was chosen as a method to express the cloned gene and synthesize NA protein. The SV40 virus has a small, double-stranded DNA genome making it easy to manipulate. The genome organization and control of gene expression of this virus is well understood and relatively large amounts of protein are made from the SV40 early and late promoters. To

Figure 6. The catalytic mechanism of lysozyme. A hexasaccharide is bound into subsites A through F of the enzyme, in this example, poly-N-acetylglucosamine (NAG). In step 1, the un-ionized carboxyl of Glu 35 donates a proton to the glycosidic bond, thereby breaking this bond. The disaccharide in subsites E and F can diffuse away. At the same time, a carboxonium ion intermediate forms at $C_1$ of the residue in subsite D, which is stabilized by the ionized carboxyl of Asp 52. In step 2, a water molecule contributes a proton to the carboxyl of Glu 35 and a hydroxyl to the carboxonium ion. After the product diffuses from the active site, the enzyme is ready for another round of catalysis. Adapted from Stryer, 1981.



study the amino acids important in the activity of the NA enzyme, site-specific mutations were made in residues of interest by subcloning the gene into an M13 vector and using synthetic mismatched oligonucleotides to direct the mutation. This mutation procedure is a powerful tool since it allows the precise mutation of a particular nucleotide or group of nucleotides in the gene of interest making any amino acid substitution desireable. It has been used successfully to study details of a number of enzyme molecules, including trypsin (Craik et al., 1985), tyrosyl tRNA synthetase (Fersht et al., 1985), carboxypeptidase A (Gardell et al., 1985) and dihydrofolate reductase (Howell et al., 1986).

In the experiments described in this dissertation, the effect of the mutation on NA enzyme activity was determined by expressing the mutant gene from the SV40 vector, assaying the infected cells for the presence of active NA, and comparing these results with the wild-type NA. Analyzing the results of a large number of mutants has led to a better understanding of the active site and a model for the mechanism of the enzyme.

This dissertation can be divided into several parts. The construction and sequencing of a full-length cDNA clone of the A/Tokyo/3/67 NA gene is described in the first of three papers that follow. This paper also includes the sequences of a number of monoclonal variants to the Tokyo/67 NA and the conclusions based on this sequence data. The second paper describes the construction and use of several SV40 late replacement vectors for expressing active NA protein from the cloned cDNA. The experimental details of the mutagenesis are presented in the second paper, and the rationale for each mutation is presented in the second and third paper. Mutagenesis was followed by expressing the mutant genes from the SV40 vector and assaying the proteins for proper expression and enzyme activity. Some preliminary results are described in paper two, and the bulk of the data and conclusions from this work are presented in the third paper. A summary and discussion of the results follows the third paper.

References

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

Air, G. M., Els, M. C., Brown, L. E., Laver, W. G., and Webster, R. G. (1985a). Location of antigenic sites on the three-dimensional structure of the influenza N2 virus neuraminidase. Virology 145, 237–248.

Air, G. M., Ritchie, L. R., Laver, W. G., and Colman, P. M. (1985b). Gene and protein sequence of an influenza neuraminidase with hemagglutinin activity. Virology 145, 117-122.

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

Basak, S., Tomana, M., and Compans, R. W. (1985). Sialic acid is incorporated into influenza hemagglutinin glycoproteins in the absence of viral neuraminidase. Virus Res. 2, 61-68.

Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. (1967). Crystallographic studies of the activity of the hen egg-white lysozyme. Proc. Roy. Soc. 167, 378-388.

Blobel, G. (1980). Intracellular protein topogenesis. Proc. Nat'l. Acad. Sci. USA 77, 1496-1500.

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

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

Breuning, A., and Scholtissek, C. (1986). A reassortant between influenza A viruses (H7N2) synthesizing an enzymatically inactive neuraminidase at 40° which is not incorporated into infectious particles. Virology 150, 65-74.

Burnet, F. M. (1948). Mucins and mucoids in relation to influenza virus action. IV. Inhibition by purified mucoid of infection and hemagglutinin with the virus strain WSE. Aust. J. Exp. Biol. Med. Sci. 26, 381-387.

Burnet, F. M., McCrea, J. F., and Anderson, S. G. (1947). Mucin as substrate of enzyme action by viruses of the mumps influenza group. Nature (London) 160, 404-405.

Cantz, M. (1982). Sialidoses. In: "Sialic Acids" (R. Schauer, ed.), pp. 307-320. Springer-Verlag, New York.

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

Colman, P. M. (1986). In: "The Influenza Viruses". (R. M. Krug, ed.). Plenum Press, New York.

Colman, P. M., and Ward, C. W. (1985). Structure and diversity of the influenza virus neuraminidase. Curr. Topics Microbiol. and Immunol. 114, 177-255.

Colman, P. M., Varghese, J. N., and Laver, W. G. (1983). Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature (London) 303, 41-44.

Compans, R. W., Dimmock, N. J., and Meier-Ewert, H. (1969). Effect of antibody to neuraminidase on the maturation and hemagglutinating activity of an influenza A2 virus. J. Virol. 4, 528-534.

Compans, R. W., Klenk, H. D., Caliguiri, L. A., and Choppin, P. W. (1970). Influenza virus proteins. I. Analysis of polypeptides of the virion and identification of spike glycoproteins. Virology 42, 880-889.

Compans, R. W., Content, J., and Deusberg, P. H. (1972). Structure of the ribonucleoprotein of influenza virus. J. Virol. 10, 795-800.

Compans, R. W. (1973). Influenza virus proteins. II. Association with components of the cytoplasm. Virology 51, 56-70.

Corfield, A. P., Higa, H., Paulson, J. C., and Schauer, R. (1983). The specificity of viral and bacterial sialidases for α(2-3) and α(2-6) linked sialic acids in glycoproteins. Biochim. Biophys. Acta. 744, 121-126.

Corfield, A. P., Wember, M., Schauer, R., and Rott, R. (1982). The specificity of viral sialidases. The use of oligosaccharide substrates to probe enzyme characteristics and strain specific differences. Eur. J. Biochem. 124, 521-525.

Craik, C. S., Largman, C., Fletcher, T., Roczniak, S., Barr, P. J., Fletterick, R., and Rutter, W .J. (1985). Redesigning trypsin: alteration of substrate specificity. Science 228, 291-297.

Dale, B., Brown, R., Miller, J., White, R. T., Air, G. M., and Cordell, B. Nucleotide and deduced amino acid sequence of the influenza neuraminidase genes of two equine serotypes. Virology, in press.

Daniels, R. S., Douglas, A. R., Gonsalves-Scarano, F., Palu, G., Skehel, J. J., Brown, E., Knossow, M., Wilson, I. A., and Wiley, D. C. (1983a). Antigenic structure of influenza virus hemagglutinin. In: "The Origin of Pandemic Influenza Viruses". (W. G. Laver, ed.), pp. 9-18. Elsevier, New York.

Daniels, R. S., Downie, J. C., Hay, A. J., Knossow, M., Skehel, J. J., Wang, M. L., and Wiley, D. C. (1985). Fusion mutants of the influenza virus hemagglutinin glycoprotein. Cell 40, 431-439.

Davis, A. R., Bos, T. J., and Nayak, D. P. (1983). Active influenza virus neuraminidase is expressed in monkey cells from cDNA cloned in simian virus 40 vectors. Proc. Nat'l. Acad. Sci. USA 80, 3976-3980.

Desselberger, U., Racaniello, V. R., Zarza, J. J., and Palese, P. (1980). The 3' and 5'-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. Gene 8, 315-328.

Doms, R. W., Gething, M. J., Henneberry, J., White, J., and Helenius, A. (1986). Variant influenza virus hemagglutinin that induces fusion at elevated pH. J. Virol. 57, 603-613.

Drzeniek, R. (1972). Viral and bacterial neuraminidases. Curr. Topics Microbiol. Immunol. 59, 35-74.

Drzeniek, R. (1973). Substrate specificity of neuraminidases. Histochem. J. 5, 271-290.

Elder, K. T., Bye, J. M., Skehel, J. J., Waterfield, M. D., and Smith, A. E. (1979). In vitro synthesis, glycosylation and membrane insertion of influenza virus hemagglutinin. Virology 95, 343-350.

Fang, R., Min Jou, W., Huylebroek, D., Devos, R., and Fiers, W. (1981). Complete structure of A/duck/Ukraine/63 influenza hemagglutinin gene: animal virus as progenitor of human H3 Hong Kong 1968 influenza hemagglutinin. Cell 25, 315-323.

Fersht, A. R., Shi, J-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., and Winter, G. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. Nature (London) 314, 235-238.

Fields, S., Winter, G., and Brownlee, G. G. (1981). Structure of the neuraminidase in human influenza virus A/PR/8/34. Nature (London) 290, 213-217.

Gardell, S. J., Craik, C. S., Hilvert, D., Urdea, M. S., and Rutter, W. J. (1985). Site-directed mutagenesis shows that tyrosine 248 of carboxypeptidase A does not play a crucial role in catalysis. Nature (London) 317, 551-555.

Gething, M.-J., and Sambrook, J. (1981). Cell-surface expression of influenza hemagglutinin from a cloned DNA copy of the RNA gene. Nature (London) 293, 620-625.

Gething, M-J. and Sambrook, J. (1982). Construction of influenza hemagglutinin genes that code for intracellular and secreted forms of the protein. Nature (London) 300, 598-603.

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

Griffin, J. A., and Compans, R. W. (1979). Effect of cytochalasin B on the maturation of enveloped viruses. J. Exp. Med. 150, 379-391.

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

Hirst, G. K. (1942). Adsorption of influenza hemagglutinins and virus by red blood cells. J. Exp. Med. 76, 195-209.

Howell, E. E., Villafranca, J. E., Warren, M. S., Oatley, S. J., and Kraut, J. (1986). Functional role of aspartic acid-27 in dihydrofolate reductase revealed by mutagenesis. Science 231, 1123-1128.

Huang, R. T. C., Wahn, K., Klenk, H.-D., and Rott, R. (1980). Fusion between cell membranes and liposomes containing the glycoproteins of influenza virus. Virology 104, 294-302.

Inglis, S. C., Barrett, T., Brown, C. M., and Almond, J. W. (1979). The smallest genome RNA segment of influenza virus contains two genes that may overlap. Proc. Nat'l. Acad. Sci. USA 76, 3790-3794.

Kawakami, K., and Ishihama, H. (1983). RNA polymerase of influenza virus. III. Isolation of RNA polymerase-RNA complexes from influenza PR8. J. Biochem. (Tokyo) 93, 989-996.

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

Klenk, H. D., Wollert, W., Rott, R., and Scholtissek, C. (1974). Association of influenza virus proteins with cytoplasmic fractions. Virology 57, 28-41.

Koennecke, I., Boschek, C. B., and Scholtissek, C. (1981). Isolation and properties of a temperature-sensitive mutant (ts412) of an influenza A virus recombinant with a ts lesion in the gene coding for the nonstructural protein. Virology 110, 16-25.

Krug, R. M. (1983). Transcription and replication of influenza viruses. In: "Genetics of Influenza Viruses." (P. Palese and D. W. Kingsbury, eds.) pp. 70-98. Springer-Verlag, Wein/New York.

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

Lamb, R. A. (1983). The influenza virus RNA segments and their encoded proteins. In: "Genetics of Influenza Viruses." (P. Palese and D. W. Kingsbury, eds.) pp. 21-69. Springer-Verlag, Wein/New York.

Lamb, R. A., and Choppin, P. W. (1979). Segment 8 of the influenza virus genome is unique in coding for 2 polypeptides. Proc. Nat'l Acad. Sci. USA 76, 4908-4912.

Lamb, R. A., and Choppin, P. W. (1981). Identification of a second protein (M2) encoded by RNA segment 7 of influenza virus. Virology 112, 729-737.

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

Lamb, R. A., Zebedee, S. L., and Richardson, C. D. (1985). Influenza virus M2 protein is an integral membrane protein expressed on the infected-cell surface. Cell 40, 627-633.

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

Laver, W. G., Air, G. M., Webster, R. G., Gerhard, W., Ward, C. W., and Dopheide, T. A. A. (1979). Antigenic drift in type A influenza virus: Sequence differences in the

hemagglutinin of Hong Kong (H3N2) variants selected with monoclonal hybridoma antibodies. Virology 98, 226-237.

Laver, W. G., Air, G. M., Webster, R. G. and Markoff, L. J. (1982). Amino acid sequence changes in antigenic variants of type A influenza virus N2 neuraminidase. Virology 122, 450-460.

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

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

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

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

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

Lentz, M. R., Air, G. M., Laver, W. G., and Webster, R. G. (1984). Sequence of the neuraminidase gene of influenza virus A/Tokyo/3/67 and previously uncharacterized monoclonal variants. Virology 135, 257-265.

Mahy, B. W. J., Barrett. T., Briedis, D. J., Brownson, J. M., and Wolstenholme, A. J. (1980). Influence of the host cell on influenza virus replication. Phil. Trans. R. Soc. (London) B 228, 349-357.

Markoff, L., Linn, B. -C., Sveda, M. M., and Lai, C. -J. (1984). Glycosylation and surface expression of the influenza neuraminidase requires the N-terminal hydrophobic region. Mol. Cell. Biol. 4, 8-16.

McCauley, J. W., Bye, J., Elder, K., Gething, M. J., Skehel, J. J., Smith, A., and Waterfield, M. D. (1979). Influenza virus hemagglutinin signal sequence. FEBS Lett. 108, 422-428.

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

Mountford, C. E., Grossman, G., Holmes, K. T., O'Sullivan, W. J., Hampson, A. W., Raison, R. L., and Webster, R. G. (1982). Effect of monoclonal antineuraminidase antibodies on the kinetc behavior of influenza virus neuraminidase. Molec. Immunol. 19, 811-816.

Mueller, O. T., Henry, W. M., Haley, L. L., Byers, M. G., Eddy, R. L., and Shows, T. B. (1986). Sialidosis and galactosialidosis: Chromosomal assignment of two genes associated with neuraminidase-deficiency disorders. Proc. Nat'l. Acad. Sci. USA 83, 1817-1821.

Murphy, B. R., and Webster, R. G. (1985). Influenza viruses. In. "Virology" (B. N. Fields, ed.), pp. 1179-1239. Raven, New York.

Murti, K. G., and Webster, R. G. (1986). Distribution of hemagglutinin and neuraminidase on influenza virions as revealed by immunoelectron microscopy. Virology 149, 36-43.

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

Newton, S. E., Air, G. M., Webster, R. G., and Laver, W. G. (1983). Sequence of the hemagglutinin gene of influenza virus A/Memphis/1/71 and previously uncharacterized monoclonal antibody-derived variants. Virology 128, 495-501.

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

Palese, P. W., and Compans, R. W. (1976). Inhibition of influenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoro-acetyl-neuraminic acid (FANA): mechanism of action. J. Gen. Virol. 33, 159-163.

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

Palese, P., Tobita, K., Ueda, M., and Compans, R. W. (1974). Characterization of temperature-sensitive influenza virus mutants defective in neuraminidase. Virology 61, 397-410.

Patel, V. (1978). Degredation of glycoproteins. In: "The Glycoconjugates, Vol. II: Mammalian Glycoproteins, Glycolipids, and Proteoglycans." (M. I. Horowitz and W. Pigman, eds.), pp. 186-229. Academic Press, New York.

Paulson, J. C., Weinstein, J., Dorland, L., Van Halbeek, H., and Vliegenthart, J. F. G. (1982). Newcastle disease virus contains a linkage-specific glycoprotein sialidase. J. Biol. Chem. 257, 12734-12738.

Phillips, D. C. (1967) The hen egg-white lysozyme molecule. Proc. Nat'l. Acad. Sci. USA 57, 484-495.

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

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

Privalsky, M. L., and Penhoet, E. E. (1978). Influenza virus proteins: identity, synthesis and modification analyzed by two-dimensional gel electrophoresis. Proc. Nat'l. Acad. Sci. USA. 75, 3625-3629.

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

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

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

Rogers, G. N., Daniels, R. S., Skehel, J. J., Wiley, D. C., Wang, X., Higa, H. H. and Paulson, J. C. (1985). Host-mediated selection of influenza receptor variants. Sialic acid α(2,6) Gal-specific clones of A/duck/Ukraine/1/63 revert to sialic acid α(2,3) Gal-specific wild type in ovo. J. Biol. Chem. 260, 7362-7367.

Rogers, G. N., Paulson, J. C., Daniels, R. S., Skehel, J. J., Wilson, I. A. and Wiley, D. C. (1983). Single amino acid substitutions in influenza hemagglutinin change receptor-binding specificity. Nature (London) 304, 76-78.

Schauer, R. (1985). Sialic acids and their role as biological masks. TIBS 10, 357-360.

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

Scholtissek, C., van Hoyningen, V., and Rott, R. (1978). Genetic relatedness between the new 1977 epidemic strains (H1N1) of influenza and human influenza strains isolated between 1947 and 1957. Virology 89, 613-617.

Schulman, J. L., and Palese, P. (1977). Virulence factors of influenza A viruses: WSN virus neuraminidase required for plaque production in MDBK cells. J. Virol. 24, 170-176.

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

Seto, J. T., and Rott, R. (1966). Functional significance of sialidase during influenza virus multiplication. Virology 30, 731-737.

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

Stryer, L. (1981). "Biochemistry." Second Edition. W. H. Freeman and Co., San Francisco.

Tulloch, P. A., Colman, P. M., Davis, P. C., Laver, W. G., Webster, R. G., and Air, G. M. (1986). Electron and X-ray diffraction studies of influenza neuraminidase complexed with monoclonal antibodies. J. Mol. Biol., in press.

Ulmanen, I., Broni, B. A., and Krug, R. M. (1981). Role of two influenza virus core P proteins in recognizing cap 1 structure ($m^7GpppNm$) on RNAs. Proc. Nat'l. Acad. Sci. USA 78, 7355-7359.

Varghese, J. N., Laver, W. G. and Colman, P. M. (1983). Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9Å resolution. Nature (London) 303, 35-40.

Ward, C. W. (1981). Structure of the influenza virus hemagglutinin. Curr. Topics Microbiol. Immunol. 94/95, 1-74.

Ward, C. W. and Dopheide, T. A. (1981). The Hong Kong hemagglutinin. Structural relationships between the human (H3) hemagglutinin and the hemagglutinin from the putative progenitor strain A/duck/Ukraine/1/63 (HAv7). In. "Genetic Variation Among Influenza Viruses" (D. P. Nayak, and C. F. Fox, eds) pp. 323-340. Academic Press, New York.

Webster, R. G., Brown, L. E. and Laver, W. G. (1984). Antigenic and biological characterization of influenza virus neuraminidase (N2) with monoclonal antibodies. Virology 135, 30-42.

Webster, R. G., Hinshaw, V. S., Bean, W. J., van Wyke, K. L., Geraci, J. R., St. Aubin, D. J. and Petursson, G. (1981). Characterization of an influenza A virus from seals. Virology 113, 712-724.

Webster, R. G., Hinshaw, V. S., and Laver, W. G. (1982). Selection and analysis of antigenic variants of the neuraminidase of N2 influenza viruses with monoclonal antibodies. Virology 117, 93-104.

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

Wilson, I. A., Skehel, J. J., and Wiley, D. C. (1981). Structure of the hemagglutinin membrane glycoprotein at 3Å resolution. Nature (London) 289, 366-373

Winter, G., and Fields, S. (1982). Nucleotide sequence of human influenza A/PR/8/34 segment 2. Nucleic Acids Res. 10, 2135-2143.

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

# SEQUENCE OF THE NEURAMINIDASE GENE OF INFLUENZA VIRUS A/TOKYO/3/67 AND PREVIOUSLY UNCHARACTERIZED MONOCLONAL VARIANTS

M. R. Lentz,[1] G. M. Air, [1] W. G. Laver,[2] and R. G. Webster[3]

[1]Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294;

[2]John Curtin School of Medical Research, Australian National University, P.O. Box 304, Canberra, A.C.T. 2601, Australia; and [3]St. Jude Childrens Research Hospital, 332 North Lauderdale, P.O. Box 318, Memphis, Tennessee 38101

Virology 135, 257-265 (1984)

## SUMMARY

A full-length cDNA copy of the neuraminidase (NA) gene of influenza strain A/Tokyo/3/67 was cloned into the plasmid pBR322, and the nucleotide sequence of the gene was determined. In addition, the sequence changes in six variants of A/Tokyo/3/67 selected with various monoclonal antibodies (Ab) to the NA were determined by dideoxy sequencing of the viral RNA. In five of the monoclonal variants, a single change occurred, resulting in an amino acid substitution at residue 344. Arginine in the parent virus changed to every amino acid possible with a single nucleotide change. In another variant, arginine at position 253 changed to serine, a change that also occurred in field strains. All variants so far sequenced that were selected by monoclonal Ab to Tokyo/3/67 virus changed at position 344, except one which changed at residue 368. Both of these residues are in clusters of residues that vary considerably in field strains, the clusters being 344-347 and 368-370. Analysis of the three-dimensional crystal structure of the NA of A/Tokyo/3/67 shows that these clusters are directly adjacent on the protein, and likely comprise a single antigenic site. A total of three or four antigenic sites have been proposed for the NA protein, based on antigenic mapping with monoclonal Ab [R. G. Webster, V. S. Hinshaw, and W. G. Laver (1982) Virology 117, 93-104]. Variants selected by Ab to Tokyo/67 NA all change in this single antigenic site, whereas variants selected by Ab to other strains change in other regions. It is possible that, although there may be three of four antigenic sites on the NA molecule, there may be a single, dominant antigenic site for each strain.

# INTRODUCTION

Neuraminidase (NA) is one of the two surface glycoproteins of influenza A and B viruses, the other being the hemagglutinin (HA). Although the precise role of the NA in viral infection is not known, it is believed to facilitate movement of the virus into and out of host cells by cleaving the HA receptor and by preventing aggregation of progeny viruses (Seto and Rott, 1966; Webster and Laver, 1967; Palese et al., 1974). The NA molecule is synthesized as a single polypeptide chain, with no post-translational modification occurring other than glycosylation (Blok et al., 1982; Fields et al., 1981). The mature protein is a tetramer of these polypeptides. Unlike the HA, which is anchored by its C-terminus (Skehel and Waterfield, 1975; Dopheide and Ward, 1981), the NA molecule is attached to the viral envelope by a hydrophobic region near its amino-terminal end (Blok et al., 1982; Fields et al., 1981). The biologically active part of the molecule is a square, box-like head attached to the membrane by a long, narrow stalk. Intact NA heads can be isolated by treating the virus with Pronase, which cleaves the NA near the junction of the head and the stalk (Laver et al., 1982). These heads retain all the enzymatic and antigenic properties of the intact NA.

Like the HA, NA undergoes antigenic variation, allowing successive viruses to escape the hosts' immune defenses. Variation occurs by periodic antigenic shifts, resulting in antigenically unrelated subtypes, and by antigenic drift where minor antigenic differences appear. Shifts occur when a virus carrying one antigenic subtype replaces a strain carrying a different subtype. This may occur by a recombination event in a doubly infected cell. Antigenic shift occurred in the NA in 1957 when the H1N1 strain was replaced by H2N2 and the NA shifted from N1 to N2. The H3N2 subtype replaced H2N2 in 1968 when only the HA shifted subtypes. H1N1 reappeared in 1977 and has cocirculated with H3N2 since then. Antigenic drift occurs when minor antigenic changes result from an accumulation of point mutations which lead to amino acid substitutions.

Several antigenic sites have been proposed for the NA molecule. This is based on studies of antigenic variants derived by growing the virus in the presence of monoclonal antibody (Ab) to the NA. Analysis of a panel of these variants with monoclonal Ab to N2 NA suggests that at least three and possibly four nonoverlapping antigenic regions are present on the surface of the NA of influenza A/Tokyo/3/67 (H2N2) (Webster et al., 1982). This is supported by regions of high variability among field strains during antigenic drift. The X-ray crystal structure of the NA of A/Tokyo/3/67 has recently been determined to 3-Å resolution (Varghese et al., 1983; Colman et al., 1983), and this should allow a more detailed analysis of these antigenic regions.

Antigenic variants of the NA of influenza A/Tokyo/3/67 were selected after a single passage of the virus in the presence of monoclonal Ab to the NA. Thirteen variants, which to not bind the Ab used for their selection, were isolated in all, at a rate of approximately 1 in $10^5$ (Webster et al., 1982). Amino acid substitutions for several of the variants were determined by amino acid compositions of tryptic peptides, and have been reported earlier (Laver et al., 1982). We have now deduced the amino acid changes for six more variants from nucleotide substitutions determined by dideoxy chain termination sequencing of the vRNA from these variants. The sequence changes of the variants and an analysis of all variants is reported in this paper. We also present the complete nucleotide sequence of the NA of the A/Tokyo/3/67 parent virus.

## MATERIALS AND METHODS

Viruses. Recombinant viruses containing the HA of A/NWS/33 (H1N1) and NA from A/Tokyo/3/67 (H2N2) or its variants were grown in the allantoic cavity of 11-day-old embryonated chicken eggs. Virus was purified by adsorption to and elution from chicken erythrocytes, followed by density gradient centrifugation in 10-40% sucrose (Laver, 1969).

Monoclonal Ab. Hybrid cell lines producing Ab to the NA of N2 subtypes were selected as described (Köhler and Milstein, 1976).

Selection of Antigenic Variants. Antigenic variants were selected as previously described (Webster et al., 1982). Briefly, monoclonal hybridoma Ab plus cloned parent virus were incubated together for 30 min at 20 °. The mixture was inoculated into 11-day-old chick embryos. The viruses that grew in the presence of monoclonal Ab were harvested and "cloned" twice at limiting dilution in 11-day-old embryonated hen's eggs. The term "monoclonal variant" will be used to describe variants selected with monoclonal Ab. The frequency of antigenic variants in the cloned preparation of A/Tokyo/3/67 was determined as described (Yewdell et al., 1979; Laver et al., 1979).

Cloning of the A/Tokyo/3/67 Neuraminidase Gene. The procedure used to obtain a full-length cDNA clone of the NA gene of A/Tokyo/3/67 was similar to that described previously (Newton et al., 1983). A synthetic oligonucleotide, 5'-dAGCA$_4$GCAG$_2$, was used to prime 25 μg of total vRNA by boiling for 1 min and quick chilling in ice water. A reaction mix containing final concentrations of 50 mM Tris, pH 8.0, 50 mM KCl, 5 mM $MgCl_2$, 10 mM DTT, 20 μCi [α-$^{32}$P]dATP, 0.5 mM dATP, 1 mM dCTP, 1 mM dGTP, and 1 mM dTTP was added to the RNA. AMV reverse transcriptase (75 U) was added to a final volume of 50 μl, and cDNA synthesis was carried out for 1 hr at 37°.

The reaction was terminated by boiling for 1 min, quick chilling, and treating the mix with 10 μg RNase for 20 min at 37°. cDNA was denatured by boiling for 3 min in formamide which contained bromphenol blue, xylene cyanol, and 50 mM EDTA. The cDNA bands were separated on a 3% acrylamide gel containing 7 M urea and Tris/borate/EDTA (TBE) buffer. Bands were visualized by autoradiography. The sixth largest band, corresponding to NA, was cut from the gel, and DNA was eluted from the gel electrophoretically. Second strand synthesis was performed under identical conditions except that 1 mM dATP was used and no radionucleotides were included. The primer used, 5'-dAGTAGA$_3$GA$_2$G, is complementary to the 3' end of the single-stranded cDNA. After second strand synthesis was completed, 10 U of the Klenow fragment of DNA polymerase I was added to ensure that the double-stranded segments had blunt ends. The double-stranded cDNA was purified by phenol-ether extraction and ethanol precipitation.

Synthetic oligonucleotide linkers containing the recognition sequence of the restriction endonuclease Sal I were phosphorylated and ligated to the cDNA. After digestion with Sal I, the cDNA was ligated to 50 ng of Sal I-cut, phosphatase-treated plasmid pBR322 and transformed into competent Escherichia coli RR1 cells.

Positive clones were determined by ampicillin resistance, tetracycline sensitivity, ability to hybridize to an influenza-specific probe, and presence of an insert of the expected size which cut out with Sal I. A large-scale plasmid stock was prepared as described (Birnboim and Doly, 1979).

Sequencing of cDNA and vRNA. The cDNA clone was sequenced by the base-specific chemical cleavage method (Maxam and Gilbert, 1980). Restriction fragments were radioactively labeled by filling in 3'-nucleotide extensions with labeled deoxynucleoside triphosphates using the Klenow fragment of DNA polymerase I. Fragments labeled at one end only were obtained by cutting with a second enzyme within labeled fragments or by isolating single strands on strand separation gels (Szalay et al., 1977). Acrylamide sequencing gels (8, 12, or 20%; 0.4 X 200 X 400 mm) were used, containing 7 M urea and TBE buffer.

vRNA coding for the NA gene of A/Tokyo/3/67 and the antigenic variants was sequenced by the dideoxy chain-termination method (Sanger et al., 1977; Air, 1979). Synthetic oligonucleotide primers were made corresponding to regions of the gene that showed no nucleotide substitutions in known sequences, and were spaced at 200- to 250-nucleotide intervals. Primers corresponded to nucleotides 226-237, 427-441, 664-678, 907-921, and 1162-1176. The first 225 nucleotides of the Tokyo/67 NA clone was sequenced completely on both strands. Since this region corresponds to the hydrophobic-anchor and stalk-sequence coding regions, which are not included in the NA heads which have the same antigenic properties of intact NA, this region did not need to be sequenced in the antigenic variants. Therefore, no primer was included for the first 237 nucleotides.

For the earlier experiments, [$\alpha$-$^{32}$P]dATP was used as the radiolabeled nucleotide, but for primers 5 and 6, [$\alpha$-thio-$^{35}$S]dATP was used since sharper bands result and the

sequence can be read more easily (Biggin et al., 1983). Sequencing gels were the same as for Maxam and Gilbert sequencing of DNA, except that only 8% gels were used, run until either the bromphenol blue or xylene cyanol was just off the gel.

## RESULTS

Cloning and Sequencing of the A/Tokyo/3/67 Neuraminidase Gene. Three clones were obtained that were ampicillin resistant, tetracycline sensitive, and hybridized to an influenza specific probe. Of these, two had inserts of the size expected for a full-length copy of the NA gene that cut out with Sal I.

The ends of the inserts of both these clones were sequenced by cutting with Sal I, labeling the ends by filling in with the Klenow fragment of DNA polymerase I using $\alpha$-$^{32}$P-labeled nucleotides, cutting again with Hae III, and isolating the labeled fragments unique to the clone. These fragments were sequenced by the Maxam and Gilbert chemical cleavage technique, and showed that both clones had the nucleotide-linker sequence followed by the conserved 3' and 5' termini of the influenza gene segments. In addition, the 5' region of both clones in the message sense coded for the six highly conserved residues of the cytoplasmic tail of previously sequenced NA genes, indicating that both clones coded for NA. Both clones therefore appeared to contain the complete gene of Tokyo/67 NA.

Approximately 82% of clone 107 was then sequenced by the chemical method. At this point, the sequence was analyzed and compared to other known N2 sequences, namely A/RI/5$^{-}$/57 (Azad et al., 1983), A/NT/60/68 (Bentley and Brownlee, 1982), A/Udorn/72 (Markoff and Lai, 1982) and A/Victoria/75 (Van Rompuy et al., 1982). Synthetic oligonucleotide primers were made by Dr. Clayton Naeve using an Applied Biosystems oligonucleotide synthesizer, and the sequence from residue 238 of the vRNA to the end of the gene was determined by the dideoxy chain-termination procedure. The dideoxy approach using the vRNA as a template eliminated potential problems of cloning out minor

variations in the viral population, and is suitable for sequencing the NA gene of related strains as well as monoclonal laboratory variants.

The strategy for determining the complete sequence of A/Tokyo/3/67 NA is shown in Figure 1. About 70% of the sequence was determined both by Maxam and Gilbert sequencing of the clone DNA and dideoxy sequencing of the vRNA. The sequence was identical at all nucleotides obtained by both methods. The complete sequence of Tokyo/67 NA is shown in Figure 2. The predicted amino acid sequence matches that of the partial sequence previously determined from NA peptides, with the exception of residue 199 which we found to be lysine (AAA) instead of the aspartic acid reported by Ward et al (1982).

Sequence Changes in Monoclonal Variants. Thirteen variants of the Tokyo/67 NA were selected, and the antigenic properties of these variants have been described (Webster et al., 1982). The amino acid changes for six of these variants, NT 32/3 V1 (Wally), NT 32/3 V2 (Colin), NT 25/4 V2 (Jan), NT S10/1 V1 (Pot), NT 25/3 V1 (Theo), and Tx 18/1 V1 (Pie), were determined by amino acid analysis of NA as previously reported (Laver et al., 1982). We have sequenced the entire head region of the remaining seven variants, NT 16/8 V1 (Gum), NT 16/8 V2 (Pear), NT 23/9 V1 (Yew), NT 23/9 V2 (Nut), NT 23/9 V3 (Elm), Jap 113/2 V1 (Sloe), and Tx 67/1 V1 (Cot), by the dideoxy chain-termination method using the primers described above. The amino acid and nucleotide sequence changes for all variants are shown in Table 1.

All variants sequenced by the dideoxy method showed only a single nucleotide change. Five of these variants, Elm, Nut, Yew, Gum, and Pear, were selected by Ab NT 23/9 or NT 16/8 to the Tokyo/67 NA. All of these variants changed at residue 344. Arginine 344 in the parent virus changed to every amino acid possible with a single nucleotide change. Changes previously determined by peptide analysis in other variants selected by Ab to Tokyo/67 NA, Jan, Theo, Colin, Wally, and Pot all show a change of arginine 344 to isoleucine, with the exception of Pot, selected by Ab S10/1, which changed from lysine to glutamic acid at position 368.

Another variant, Sloe, selected by Ab 113/2 to A/Japan/305/57 NA, showed a change of A to T at nucleotide 778 (Figure 3), resulting in an amino acid substitution of arginine 253 to serine. The seventh variant sequenced, Cot, selected by Ab Tx 67/1 to the Texas/77 strain, showed no nucleotide substitutions, and Ab-binding assays repeated on the virus preparation confirmed that it had reverted back to wild-type during growth for RNA isolation.

## DISCUSSION

Work with monoclonal variants has been undertaken as a method of elucidating sites of antigenic importance on the NA molecule. With the recent determination of the three-dimensional structure of the NA of A/Tokyo/3/67, regions of high variability in field strains and changes in monoclonal variants can be correlated to molecular structure (Coleman et al., 1983). In this paper we have analyzed the sequence of previously uncharacterized variants by nucleotide sequencing of vRNA. These changes and those of other monoclonal variants previously determined by amino acid analysis can now be located in the three-dimensional structure of the NA protein. In addition, the complete nucleotide sequence of the NA gene of influenza A/Tokyo/3/67 was determined.

There are several clusters of residues that are highly variable in field strains of N2 NA. These clusters include residues 344-347 and 368-370. Ab binding assays indicate that these clusters are part of a single antigenic site (Webster et al., 1982), and are now known to be adjacent on the three-dimensional crystal structure. The variants sequenced, which were selected by Ab to Tokyo/67 NA, Elm, Nut, Yew, Gum, and Pear, all had amino acid substitutions at the same residue, 344. Four other Tokyo Ab-derived variants whose changes were previously determined by peptide analysis, Jan, Theo, Colin, and Wally, also change at this residue, and a tenth variant, Pot, also selected by a Tokyo/67 Ab, changed at adjacent residue 368. All of these variants were selected by Ab which fall into antigenic groups Ia or Ib, which overlap but can be distinguished in ELISA assays (Webster et al., 1982). These monoclonal variant changes therefore support the Ab-

binding and field-strain variability data, suggesting that this region of the NA molecule is an antigenically important one. As shown in Figure 4, these residues are located on the upper surface of the molecule and would clearly be accessible to an Ab molecule. It is interesting to note the change that occurred in the variant Yew. This variant changed at residue 344 from arginine to lysine, generally considered to be a very conservative substitution. Most characteristics of this position, such as charge, size, and shape of the side chain are retained, yet the ability of the Ab 23/9 to bind is completely lost. This same residue changes to every amino acid possible with a single nucleotide substitution.

Variant Pie was selected by Ab Tx 18/1 to the NA of Texas/77 strain, an Ab that falls into antigenic group II. This variant showed a sequence change of asparagine 221 to histidine, a position that does not vary in any of the sequenced field strains. No strains have been sequenced beyond 1975 isolates, however, and it is possible that in these more recent strains, residue 221 does become antigenically important. This residue is also located on the upper surface of the NA molecule (Figure 4), where an Ab would be free to bind.

The final variant we have sequenced is Sloe, which was selected by Ab Jap 113/2 (group III) to Japan/57 NA. This variant changed at residue 253 from arginine to serine, a substitution which is more difficult to explain. As can be seen in Figure 4, this residue is located on the three-dimensional structure in a small pocket on the underside on the molecule. It is unlikely that an Ab would have easy access to this part of the NA and, furthermore, the combining site would be expected to be too large to enter the pocket and reach this particular side chain. This is the only substitution in this variant, however; therefore it must be responsible for the lack of Ab binding, and the same residue has also changed in field strains. It is possible that a substitution at this position could transmit a conformational change further along the polypeptide chain and alter an Ab-binding site on the upper surface of the molecule.

The sequence data on the monclonal variants suggests that, although there may be three or more distinct antigenic sites on the NA molecule, there may be a single, dominant

site for each strain. Different variants selected by a variety of Ab to the Tokyo/67 NA all changed at a single antigenic site, suggesting that this site is dominant in the Tokyo/67 strain, whereas variants selected by Ab to different strains had changes in other regions of the molecule.

Sequencing of monoclonal variants derived from parent strain X-7F1 by a different panel of monoclonal Ab is now underway, and this data should give a clearer picture of the antigenic properties of the influenza NA and possibly of dominant antigenic sites.

ACKNOWLEDGMENTS

We are greatly indebted to Dr. Clayton Naeve, St. Jude Childrens' Research Hospital, Memphis, for synthesizing the oligonucleotides, and to Randy Owens for making the wire model. This work was supported by grants AI-18203 and AI-19084 from NIAID, and CA 13148 from NCI. M. L. is supported by NIH Training Grant AI-07150. We thank Jean Clark and Terri Shipman for expert technical assistance. This collaborative project was greatly helped by international direct-dialing telephone facilities provided by the Australian Overseas Telecommunications Commission.

Figure 1. Strategy for sequencing the A/Tokyo/3/67 NA gene. The upper set of arrows represents the sequence obtained by chemical cleavage sequencing of restriction fragments of the cDNA clone (Maxam and Gilbert, 1980). Suitable restriction sites were chosen using the sequence of A/RI/5-/57 NA (Azad et al., 1983), and later using sequence already obtained from the clone. The arrows indicate the site from which the fragment was labeled and the length of the sequence obtained. Arrows pointing left were from the viral RNA sense strand, those to the right from the message sense strand. The lower set of arrows represents the sequence obtained by dideoxy sequencing of the vRNA (Sanger et al., 1977). Suitable primer positions were determined from the partial Maxam and Gilbert sequence of the clone cDNA and known sequences of other N2 strains. Primers, represented by circles at the ends of the arrows and synthesized by Dr. Clayton Naeve on an Applied Biosystems oligonucleotide synthesizer, are TAACACCACCAT (226-237), $G_4$ACCACACTAGA (427-441), TGGTTCATGGTCTCA (664-678), AGGCTCTAATAGGCC (907-921), and TGGTTGGTCCACACC (1162-1176). Arrows represent the length of the sequence obtained from each primer. All dideoxy sequences are in the message sense. The dideoxy sequence was determined at least three separate times from each primer.

Figure 2. Nucleotide and predicted amino acid sequence of A/Tokyo/3/67 NA. The sequence was obtained by both chemical cleavage sequencing of clone cDNA (Maxam and Gilbert, 1980) and dideoxy sequencing of vRNA (Sanger et al., 1977). Synthetic primers were used for dideoxy sequencing. The predicted amino acid sequence matches that of the partial sequence previously determined from NA peptides, with the exception of residue 199 which we found to be lysine instead of the aspartic acid reported by Ward et al (1982).

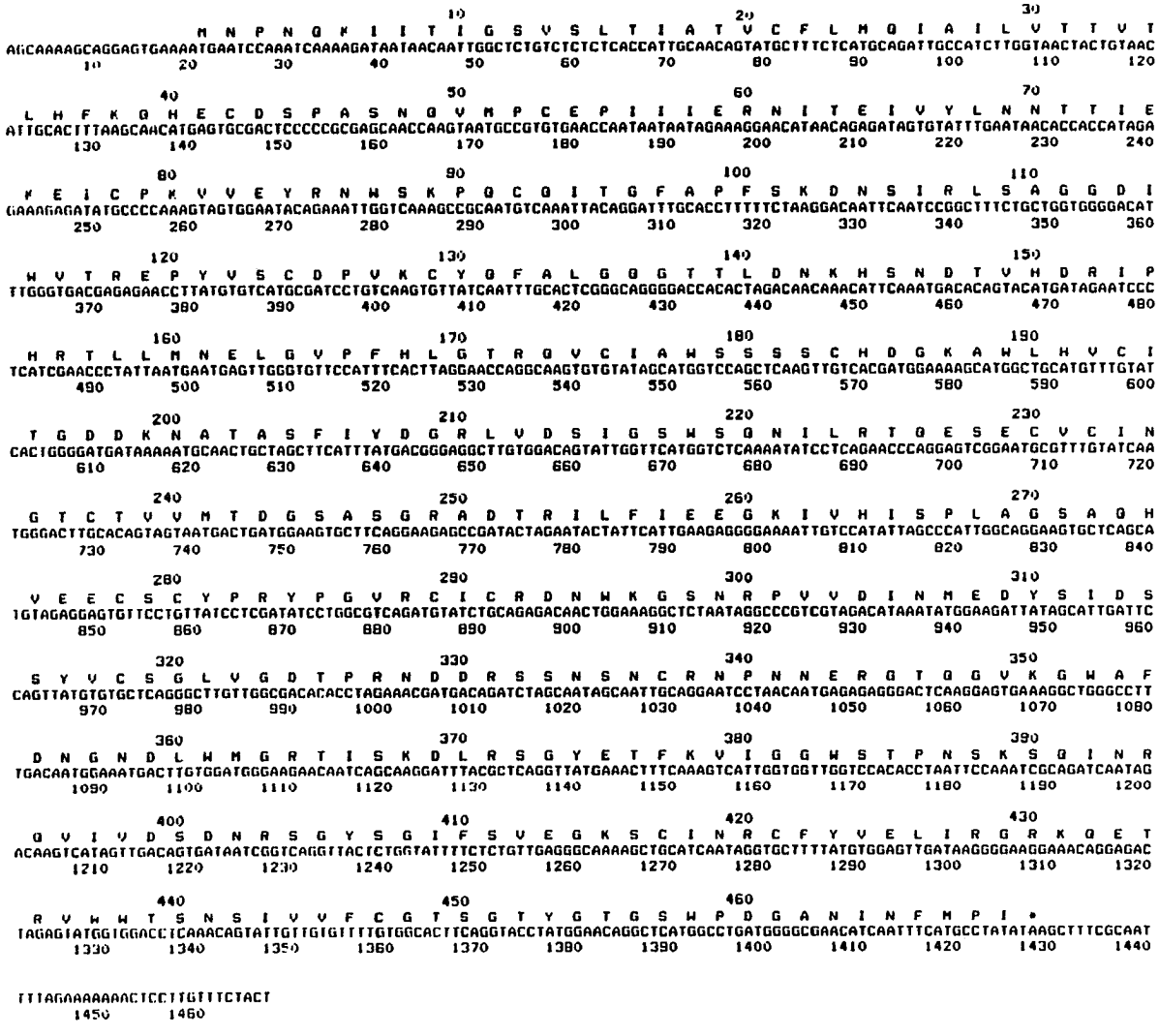


TABLE 1
AMINO ACID CHANGES IN MONOCLONAL VARIANTS
OF A/TOKYO/3/67 NEURAMINIDASE

| Variant (trivial name) | Monoclonal Ab Used for Selection | Amino Acid Change | Codon Change |
|---|---|---|---|
| V1 (Theo) | NT S25/3[a] | 344 Arg to Ile[d] | |
| V2 (Jan) | NT 25/4 | 344 Arg to Ile[d] | |
| V1 (Wally) | NT 32/3 | 344 Arg to Ile[d] | |
| V2 (Colin) | NT 32/3 | 344 Arg to Ile[d] | |
| V1 (Yew) | NT 23/9 | 344 Arg to Lys | AGA to AAA |
| V2 (Nut) | NT 23/9 | 344 Arg to Gly | AGA to GGA |
| V3 (Elm) | NT 23/9 | 344 Arg to Ile | AGA to ATA |
| V1 (Gum) | NT 16/8 | 344 Arg to Thr | AGA to ACA |
| V2 (Pear) | NT 16/8 | 344 Arg to Ser | AGA to AGT |
| V1 (Pot) | NT S10/1 | 368 Lys to Glu[d] | |
| V1 (Sloe) | Jap 113/2[b] | 253 Arg to Ser | AGA to AGT |
| V1 (Pie) | Tx 18/1[c] | 221 Asn to His[d] | |
| V1 (Cot) | Tx 67/1 | None found | |

a "NT" antibodies are made to NA of A/Tokyo/3/67

b "Jap" antibodies are made to NA of A/Japan/305/57

c "Tx" antibodies are made to NA of A/Texas/1/77

d Laver et al. (1982)

Figure 3. Sequence autoradiograph from A/Tokyo/3/67 NA and variants Cot and Sloe. The sequence is from primer 4, and the 8% gel was run until the xylene cyanol was just off the gel. The labe used was $^{32}P$, and nucleotides 768 to about 900 are shown. By running all A lanes together, then all G lanes, C lanes, and T lanes, monoclonal variant changes can be seen by merely scanning the gel by eye. On this gel, a change at position 778 of the variant Sloe can be seen where an A in the parent virus is replaced by a T. This changes the codon from AGA to AGT, and an amino acid substitution of arginine to serine occurs.

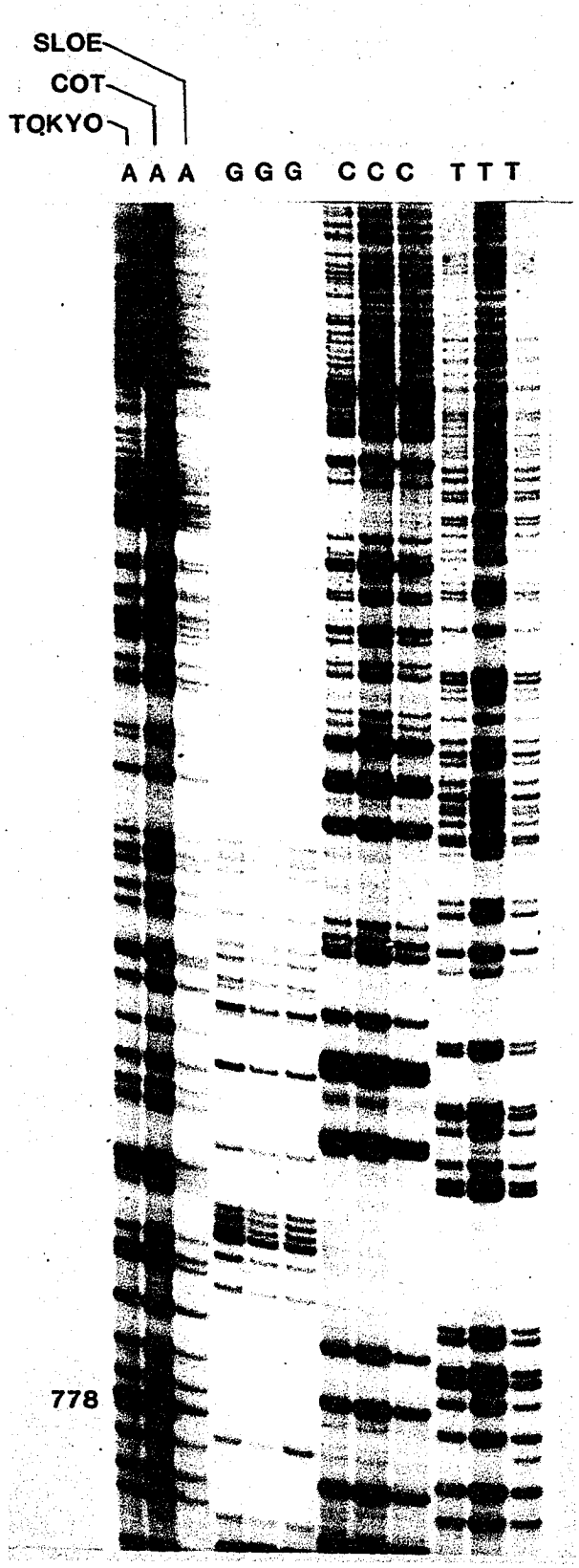
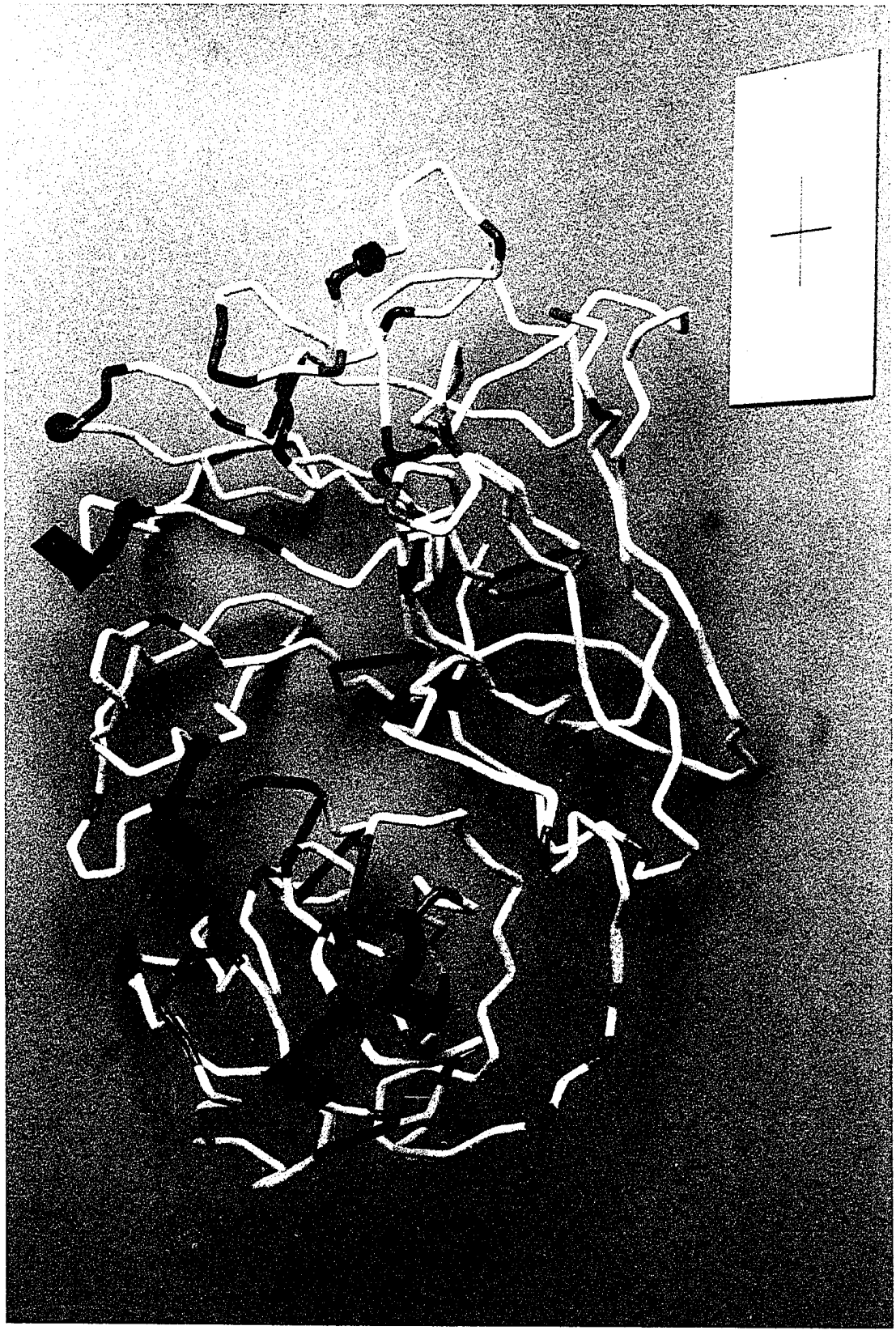


Figure 4. The three-dimensional structure of Tokyo/67 NA showing the amino acid positions of monoclonal variant changes. One subunit of the Pronase-released head is shown, and the bent wire represents the α-carbon chain tracing. Amino acid residues which have changed in the monoclonal variants are flagged. Residues 344 and 368 are adjacent and located on the upper surface near one side which contacts another subunit. Residue 221 is in a similar position on the upper surface of the molecule near the side oposite 344 and 368. Residue 253 is located in a pocket on the underside of the protein. The cross aproximately represents the four-fold axis of symmetry in the mature tetramer, and "banded" residues are glycosylated. Data for constructing the model was from Varghese et al., (1983).



## REFERENCES

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

Azad, A. A., Elleman, T. C., Laver, W. G., and Ward, C. W. (1983). Sequence changes associated with antigenic shift and drift in influenza virus neuraminidase. In "The Origin of Pandemic Influenza Viruses" (W. G. Laver, ed.), pp. 59-76. Elsevier, New York/Amsterdam.

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

Biggin, M. D., Gibson, T. J., and Hong, G. F. (1983). Buffer gradient gels and $^{35}S$ label as an aid to rapid DNA sequencing. Proc. Nat'l. Acad. Sci. USA 80, 3963-3965.

Birnboim, H. C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7, 1513-1523.

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

Colman, P. M., Varghese, J. N., and Laver, W. G. (1983). Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature (London) 303, 41-44.

Dopheide, T. A. A., and Ward, C. W. (1981). The location of the bromelain cleavage site in a Hong Kong influenza hemagglutinin. J. Gen. Virol. 52, 367-370.

Fields, S., Winter, G., and Brownlee, G. G. (1981) Structure of the neuraminidase in human influenza virus A/PR/8/34. Nature (London) 290, 213-217.

Köhler, G., and Milstein, C. (1976). Derivation of specific antibody-producing tissue culture and tumor lines by cell fusion. Eur. J. Immunol. 6, 511-519.

Laver, W. G. (1969). Purification of influenza virus. In: "Fundamental Techniques in Virology" (K. Habel and N. P. Salzman, eds.) pp. 82-86. Academic Press, New York.

Laver, W. G., Air, G. M., Webster, R. G., Gerhard, W., Ward, C. W., and Dopheide, T. A. A. (1979). Antigenic drift in type A influenza virus: Sequence differences in the hemagglutinin of Hong Kong (H3N2) variants selected with monoclonal hybridoma antibodies. Virology 98, 226-237.

Laver, W. G., Air, G. M., Webster, R. G., and Markoff, L.J. (1982). Amino acid sequence changes in antigenic variants of type A influenza virus N2 neuraminidase. Virology 122, 450-460.

Markoff, L., and Lai, C-J. (1982). Sequence of the influenza A/Udorn/72 (H3N2) virus neuraminidase gene as determined from cloned full-length DNA. Virology 119, 228-297.

Maxam, A. M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. In: "Methods in Enzymology" (L. Grossman and K. Moldave, eds.). Vol. 65, pp. 499-560. Academic Press, New York.

Newton, S. E., Air, G. M., Webster, R. G., and Laver, W. G. (1983). Sequence of the hemagglutinin gene of influenza virus A/Memphis/1/71 and previously uncharacterized monoclonal antibody-derived variants. Virology 128, 495-501.

Palese, P., Tobita, K, Ueda, M., and Compans, R. W. (1974). Characterization of temperature-sensitive influenza virus mutants defective in neuraminidase. Virology 61, 397-410.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Nat'l. Acad. Sci. USA 74, 5463-5467.

Seto, J. T., and Rott, R. (1966). Functional significance of sialidase during influenza virus multiplication. Virology 30, 731-737.

Skehel, J. J., and Waterfield, M. D. (1975). Studies on the primary structure of the influenza virus hemagglutinin. Proc. Nat'l. Acad. Sci. USA 72, 93-97.

Szalay, A. A., Grohman, K., and Sinsheimer, R. L. (1977). Separation of the complementary strands of DNA fragments on polyacrylamide gels. Nucl. Acids Res. 4, 1569-1578.

Van Rompuy, L., Min Jou, W., Huylebroeck, D., and Fiers, W. (1982). Complete nucleotide sequence of human influenza neuraminidase gene of subtype N2 (A/Victoria/3/75). J. Mol. Biol. 161, 1-11.

Varghese, J. N., Laver, W. G., and Colman, P. M. (1983). Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9Å resolution. Nature (London) 303, 35-40.

Ward, C. W., Elleman, T. C., and Azad, A. A. (1982). Amino acid sequence of the Pronase-released heads of neuraminidase subtype N2 from the Asian strain A/Tokyo/3/67 of influenza virus. Biochem. J. 207, 91-95.

Webster, R. G., Hinshaw, V. S., and Laver, W. G. (1982). Selection and analysis of antigenic variants of N2 influenza viruses with monoclonal antibodies. Virology 117, 93-104.

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

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

# LOSS OF ENZYME ACTIVITY IN A SITE-DIRECTED MUTANT OF INFLUENZA NEURAMINIDASE COMPARED TO EXPRESSED WILD-TYPE PROTEIN

Michael R. Lentz and Gillian M. Air

Department of Microbiology
University of Alabama at Birmingham
Birmingham, Alabama 35294

VIROLOGY 148, 74-83 (1986)

## SUMMARY

Full-length double-stranded DNA copies of the neuraminidase (NA) gene of influenza virus A/Tokyo/3/67 (N2) and a mutant generated in vitro by site-specific oligonucleotide-directed mutagenesis with a substitution of leucine for tryptophan at position 178 were cloned into an SV40 late replacement expression vector. Indirect immunofluorescence of cells infected with these recombinant vectors showed the presence of NA protein in the cytoplasm and on the surface of infected cells. Cells expressing the wild-type protein showed neuraminidase enzyme activity for both fetuin, a sialated glycoprotein (mol wt = 50,000) and N-acetylneuraminyl lactose, a trisaccharide (mol wt = 600). This enzyme activity was inhibited by 44 % toward N-acetylneuraminyl lactose and by 98% toward fetuin by adding anti-NA antibody before substrate. In contrast, cells expressing the mutant NA had no detectable enzyme activity for either substrate. The conserved nature of the tryptophan at position 178 in all known NA strains, its location in the substrate binding pocket in the three-dimensional structure and the lack of activity of the mutant protein indicate that this residue is essential for enzyme activity.

# INTRODUCTION

Influenza virus particles contain two glycoproteins, hemagglutinin (HA) and neuraminidase (NA) which are anchored to and project from the viral envelope. HA binds to host cell receptors and initiates penetration by fusing viral and host cell membranes. The NA enzyme catalyzes the hydrolytic cleavage of the α-ketosidic bond linking a terminal N-acetylneuraminic acid (NANA) residue to the adjacent sugar on polysaccharides (Gottschalk, 1957). NA is believed to facilitate movement of the virus into and out of host cells by cleaving the HA receptor and by preventing the aggregation of progeny viruses (Seto and Rott, 1966; Webster and Laver, 1967; Palese et al., 1974), and may also modify HA carbohydrate side chains (Basak et al., 1985) to allow activational cleavage of HA (Schulman and Palese, 1977). Both proteins function in immune recognition of the virus by the host. Like the HA, NA undergoes antigenic variation preventing the buildup of an effective immune response to later strains of the virus. The NA polypeptide undergoes no post-translational modification other than glycosylation (Blok et al., 1982) and assembles as a tetramer on the surface of the virus (Wrigley et al., 1973; Varghese et al., 1983). The NA polypeptide is attached to the viral envelope by a hydrophobic sequence near its amino terminal end (Blok et al., 1982; Fields et al., 1981). The biologically active part of the molecule is a square, box-like "head" attached to the envelope by a long, narrow stalk. Treatment of virus particles with Pronase releases the heads by cleaving the protein at the junction of the head and the stalk; the heads retain all the antigenic and enzymatic properties of the intact protein (Laver, 1978; Laver et al., 1982).

The X-ray structure of NA heads at 2.9 Å resolution shows a large pocket near the top of each monomer. When crystals are soaked in sialic acid prior to X-ray diffraction, a prominant structure can be seen protruding from the pocket, suggesting that this is the location of the substrate binding and enzyme active site (Colman et al., 1983).

A comparison of the known NA gene sequences shows scattered clusters of highly variable residues which presumably give rise to the antigenic differences in these strains. There are also fifteen charged residues and four uncharged residues which are conserved in all known strains and subtypes (Colman et al., 1983; Air et al., 1985). When located on the three-dimensional structure, these conserved residues are found to be within and surrounding the sialic acid binding pocket. Since it is expected that side chains involved in enzyme activity must be conserved, these residues may be involved in substrate binding or catalytic activity.

In order to study the antigenic and enzymatic properties of NA in more detail, we have constructed an SV40 late replacement vector for expressing NA. This vector has been used to express a wild-type NA (A/Tokyo/3/67) gene and a NA gene containing a mutation in the proposed active site. In this paper we present characteristics of the NA produced including the enzyme activity of these proteins.

## MATERIALS AND METHODS

DNAs, Viruses, and Cells. The cloning of a full-length cDNA copy of the NA gene of influenza virus A/Tokyo/3/67 has been reported (Lentz et al., 1984) and was the source of NA cDNA. The construction of the SV40 late replacement expression vector pQPS is described below. SV40 late functions were provided by helper virus dl 1055 (Pipas et al., 1979,1983), a viable, early deletion mutant generously provided by Dr. Eric Hunter (University of Alabama at Birmingham).

All recombinant plasmids were propagated in Escherichia coli strain RR1. CV-1 African green monkey kidney cells, permissive for SV40 replication, were used for expression of recombinant SV40 viruses. Cells were grown in Dulbecco's modified Eagle's medium containing 60 U/ml penicillin, 60 μg/ml streptomycin, and 10% fetal calf serum, except where NA assays were to be performed when serum-free medium was used.

Manipulation of DNAs. All manipulations of DNA (restriction endonuclease digestions, ligations, transformations, plasmid isolations, gel electrophoresis) were

performed according to standard protocols (Maniatis et al., 1982). Restriction endonucleases were purchased from Bethesda Research Laboratories (Gaithersburg, Md.) or New England Biolabs (Beverly, Mass.). T4 DNA ligase (Tait et al., 1980) and Klenow fragment of DNA polymerase I (Joyce and Grindley, 1983) were both grown and purified in our laboratory from cloned genes under the control of bacteriophage lambda promoters. Synthetic oligonucleotides were synthesized by Dr. Clayton Naeve on an Applied Biosystems DNA synthesizer.

Construction of the SV40 Recombinant Vector. Plasmid vector pQP was supplied to us by Dr. Eric Hunter. This vector was constructed by deleting the late coding region of SV40 and removing the Agno gene initiator codon (L. Perez, J. Wills, and E. Hunter, manuscript in preparation).

In order to fit our cloning experiments, pQP was further modified to accept Sal I fragments by deleting the Sal I site in the pBR322 sequence and inserting a synthetic oligonucleotide linker at the Cla I site downstream of the late promoter (Figure 1). This linker contains the Sal I recognition sequence and nucleotides to regenerate the Cla I site. The resulting vector, pQPS, was used in the experiments described. The bacterial sequences are removed by digestion with either Cla I or Bam HI. The complete wild-type Tokyo/67 NA gene cDNA fragment cut from pBR322 or mutant Tokyo/67 cDNA cut from M13 replicative form (RF) was subcloned into Sal I digested, phosphatase-treated pQPS, transformed into E. coli RR1 cells, and screened for the presence of the NA gene by colony hybridization (Maniatis et al., 1982). Restriction enzyme analysis was used to determine which clones had the NA gene in the correct orientation for expression.

Transfection and Virus Stocks. Bacterial plasmid sequences were removed from the recombinant vector and helper virus dl 1055 by Bam HI digestion. Two hundred nanograms of the SV40 vector and 75 ng helper DNA were ligated separately to recircularize at low DNA concentration (4μg/ml). Virions carrying the recombinant SV40-NA DNA were produced by DEAE dextran transfection using the method of Gething and Sambrook (1981) followed by a 4 hr incubation in medium containing 100 μM chloroquine

(Luthman and Magnusson, 1983). Plates were incubated at 37° for 7 days at which time cytopathic effects (CPE) were evident. Preparation of virus stocks was completed by freezing and thawing three times. A high titer stock was prepared by passaging the first lysate in CV-1 cells and used at a 1:10 dilution for the experiments described. This dilution produces moderate CPE 48-72 hr postinfection.

Mutagenesis. The full-length clone of the wild-type NA gene was subcloned into the Sal I site of M13mp18 and single-strand template DNA with minus-sense NA was isolated from the phage in a 40 ml culture of infected JM105 cells. The mutagenesis procedure of Zoller and Smith (1983) was followed with these modifications: (a) In addition to the mutagenic primer, a universal M13 primer and a NA specific sequencing primer, both of which hybridize upstream of the mutagenic primer, were used. This increases the number of covalently closed molecules synthesized in the reaction (Norris et al., 1983). (b) The mutagenesis reaction mix was used directly for transformation without enriching for covalently closed molecules. Usually 1/100 of the reaction mix was used per transformation. The mutagenic primer used, 5'TGTGTATAGCATTGTCCAG, is equivalent to nucleotides 540 to 558 of the NA gene in the message sense. The wild-type sequence has a TGG tryptophan codon at position 551 to 553 which is mutated to TTG leucine by this primer.

Plaques were screened for the mutation by dot blot hybridization using 50 µl of phage supernatant spotted onto nitrocellulose filters with a Scleicher and Schuell Minifold vacuum system. $^{32}P$ 5' end-labeled mutagenic oligonucleotide was used as a probe. After hybridization, filters were washed at successively higher temperatures and prospective mutants were sequenced by the dideoxy chain-termination method (Sanger et al., 1977; Biggin et al., 1983) from M13 template DNA to confirm the mutation.

Immunofluorescence of Infected Cells. Internal and surface indirect fluorescent staining of infected cells was carried out as described (Wills et al., 1984) except that 3.7% formalin in phosphate-buffered saline (PBS) was used instead of ethanol/acetic acid to fix surface stained cells. Two different antibody preparations were used: a polyclonal rabbit

antiserum which binds to both native and denatured Tokyo/67 NA, and a pool of monoclonal antibodies which only recognizes the native Tokyo/67 NA three-dimensional structure. Photographs were made on a Nikon Fluophot fluorescence microscope with Kodak Tri-X Pan film (400 ASA).

Radiolabeling, Immunoprecipitation, and Polyacrylamide Gel Electrophoresis of Proteins. Cells were labeled for 3 hr with L-[4,5-$^3$H(N)]leucine (New England Nuclear) 48-72 hr postinfection as desribed (Wills et al., 1984). Radiolabeled cells were lysed and the NA protein was immunoprecipitated (Hunter et al., 1983) using an excess of anti-Tokyo/67 NA polyclonal antisera. Immunoprecipitated proteins were separated on 8% polyacrylamide gels containing 8M urea, 0.1% SDS, and 0.1 M sodium phosphate, pH 7.2, overnight at 35 mA in a running buffer containing 0.1 M sodium phosphate, 0.1% SDS, and 0.013 M 2-mercaptoethanol. Gels were stained with Coomassie brilliant blue, treated with En$^3$Hance (Dupont Chemicals), dried, and fluorographed.

Neuraminidase Enzyme Assay. The method of Aymard-Henry et al (1973) was used. Cells from a 100 mm plate 48-72 hr after infection with the recombinant virus stock were scraped or freeze-thawed twice to remove the cells from the plate. Cells in suspension were then pelleted from the media for 10 min at 2000 rpm, resuspended in 100 µl of 0.25 mM $CaCl_2$, 0.8 mM $MgCl_2$, 0.15 M NaCl (CMS), and transferred to glass tubes. One hundred microliters of 40 mg ml$^{-1}$ fetuin or 1 mM N-acetylneuraminyl lactose in 0.2M $Na_2HPO_4$, pH 5.9, was added and the tubes were incubated at 37° for 1.5 hr. The N-acetylneuraminic acid released, chemically converted to a pink chromophore, was extracted into 4 ml of 95%butanol/5% hydrochloric acid (v/v), and the absorbance was read at 549 nm.

Inhibition of Enzyme Activity. Cells were removed from the plate and pelleted as described above, then resuspended in 50 µl anti-NA antibody (Ab) diluted 1:25 in PBS. The cell suspension was incubated at room temperature for 90 min. Cells were pelleted 2 min in a microfuge, resuspended in 100 µl CMS, and assayed as above. As a control, cells were also treated with anti-HA Ab in the same manner.

## RESULTS

Expression of Neuraminidase Protein. We constructed several different SV40 late replacement expression vectors which vary considerably in their ability to express NA protein from a cloned gene. NA expression from vector pQDS-NA, which has Sal I and Cla I linkers inserted at the Hpa II site of SV40, was very weak as determined by immunofluorescence. Expression of the Tokyo/67 matrix gene was also barely detectable. The sequence of this vector in the vicinity of the late promoter revealed two out-of-frame ATG codons between the late promoter and the initiator ATG of the NA gene, one being the agno gene initiating codon. It has been proposed that "upstream" AUGs siphon off ribosomes and prevent efficient initiation from AUG codons further downstream (Kozak, 1983), and this is likely to be the reason that this vector expresses NA at such low levels.

In order to eliminate the ATG codons upstream of the cloning site, a second SV40 expression vector, pQKS, was constructed by inserting a Sal I cloning site at the unique Kpn I site of SV40. No significant expression of NA protein could be detected by immunofluorescence in cells infected with this vector carrying the cloned NA gene. Nucleotides 294-304 of SV40, including the Kpn I recognition site are an important part of the SV40 late promoter (Brady et al., 1982) and it seems likely that the SV40 late promoter in this vector was inactivated although we did not determine transcription rates experimentally.

Vector pQPS, described above, was used for the remaining experiments. The NA initiator codon is 81 base pairs downstream of the Kpn I site of the SV40 late promoter. There are no ATG codons preceding the NA ATG in this construction. Recombinant SV40 virus stocks were prepared as described above and designated SVNAwt (SV40 vector expressing wild-type Tokyo/67 NA) and SVNA178L (SV40 vector expressing NA with a Trp to Leu mutation at residue 178). Indirect immunofluorescence of CV-1 cells infected with either virus stock using polyclonal rabbit anti-Tokyo/67 NA antisera revealed that NA protein was being expressed at high levels in these cells as shown in Figure 2. There was

no detectable difference in the expression of the wild-type and mutant proteins. Internal fluorescence of fixed cells revealed bright perinuclear staining typical of proteins in the Golgi apparatus where glycosylation occurs. Bright fluorescence was seen on the surface of unfixed cells expressing either the wild-type or mutant proteins (Figure 2) indicating that normal transport of the protein to the cell surface is unaffected by the mutation. Uninfected cells or cells infected with recombinant stock VH (late replacement vector with no inserted gene) showed only background fluorescence.

Radiolabeled wild-type and mutant NA proteins, immunoprecipitated from infected cells, comigrated on polyacrylamide gels along with the NA protein from whole disrupted $NWS_{HA}/Tokyo_{NA}$ virions (data not shown). This is further evidence that the NA proteins expressed from the SV40 vector are being properly synthesized, glycosylated and transported. Immunofluorescence and immunoprecipitation results indicate that the two proteins are synthesized at nearly identical levels in infected cells.

Enzyme Activity of the Neuraminidase Gene Products. NA cleavage of NANA from either fetuin, a glycoprotein of mol wt 50,000 or N-acetylneuraminyl lactose, a trisaccharide of mol wt 600 was detected colorimetrically (Aymard-Henry et al., 1973). When cells from a 100-mm plate infected with the SVNAwt virus stock were assayed 48-72 hr post infection, a bright pink color resulted with an A549 ranging from 0.3 to 0.5 (Figure 3). Calculations based on enzyme activity demonstrate that approximately 100,000 molecules of active NA are present per infected cell or approximately 20-25 ng per 100-mm plate of infected cells. This was determined by comparing the level of activity of infected cells to a known quantity of purified Tokyo/67 NA heads (mol wt = 200,000). Approximately $5 \times 10^5$ cells per plate were expressing NA, determined by counting immunofluorescent cells on a duplicate plate.

The level of activity was the same whether cells were freeze-thawed to remove them from the plate or a washed monolayer of intact cells was gently scraped from the plate and assayed. When uninfected cells or cells infected with the SVV or SVNA178L virus stocks were treated in the same way, no enzyme activity could be detected (Figure 3).

NA activity on SVNAwt-infected cells was inhibited by treating cells with anti-Tokyo/67 NA Ab prior to adding substrate as described above. As shown in Figure 3, activity on infected cells was reduced by an average of 44% toward N-acetylneuraminyl lactose and by 98% toward fetuin. This is nearly identical to the values obtained when activity on influenza virus particles is inhibited in the same manner with the same antiserum (data not shown), and similar to results obtained by Rafelson et al., (1983) for Japan/305/57 and PR/8/34 virus particles. Anti-HA Ab has no effect on the enzyme activity for either substrate (data not shown). Inhibition by anti-NA Ab provides further evidence that the enzyme activity measured is a result of the expression of the cloned influenza NA gene.

To test for possible conformational differences between mutant and wild-type NAs, internal and surface fluorescence was performed using a pool of monoclonal antibodies. In ELISA tests, this monoclone pool has high titers against whole influenza virus but does not react with boiled virus or virus treated with detergent (Lentz and Webster, unpublished results). This pool of monoclonal antibodies therefore appears to react exclusively with NA in the native three-dimensional conformation, whereas the rabbit antiserum recognized both native and denatured NA. Internal fluorescence, which was bright with the rabbit antiserum, could not be detected when the monoclonal antibodies were used in cells infected with SVNAwt or SVNA178L, suggesting that the fixing procedure partially denatures the proteins or that the proteins in the cell have not yet folded into the native structure. The surface of unfixed cells infected with SVNAwt or SVNA178L stained brightly with the monoclone pool and were indistinguishable from each other indicating that the mutant NA structure is the same as wild-type at the level of detection of this experiment.

## DISCUSSION

We have cloned full-length copies of wild-type and mutant influenza A/Tokyo/3/67 neuraminidase genes into an SV40 late replacement vector and used this system the study the expression and enzyme activity of the proteins produced. Indirect immunofluorescence

of fixed and unfixed cells revealed the presence of NA protein in the cytoplasm and on the surface of infected cells. There was no difference in the internal of surface fluorescent staining of wild-type and mutant proteins suggesting that the mutation has no effect on the expression, transport or assembly of the mutant NA. Immunoprecipitation of NA from infected cells showed that there was no difference in the migration of wild-type and mutant protein and both comigrated with NA from influenza virions.

The properties of wild-type and mutant NA appeared identical in all respects except enzyme activity. Cells infected with the SVNAwt vector consistently showed high levels of NA enzyme activity by the standard assay, whereas SVNA178L infected cells had no detectable levels of activity. The wild-type enzyme was active for two substrates used, N-acetylneuraminyl lactose and fetuin. Other laboratories have expressed NA from cloned genes in SV40 vectors, and although they observed enzyme activity for Umbelliferone-NANA (Markoff et al., 1984) and α-2,3[$^3$H]sialyl-lactitol (Davis et al., 1983), no estimations were made of the amount of NA present. Both of these are low molecular weight substrates. Because of their small size, these substrates can reach the enzyme active site with relatively little steric hindrance and therefore do not accurately reflect the natural situation where NA encounters cellular or viral glycoproteins as substrates (Markoff et al., 1984). Our detection of high levels of NA activity toward fetuin, a glycoprotein of mol wt 50,000 suggests that the NA expressed from the SV40 vector may be forming a molecule identical to that in a natural influenza infection, although since it is not known if the monomer is active, we do not know if the expressed NA is tetrameric.

Bachmeyer (1972) showed that chemical modification of tryptophan residues of NA completely inactivated the enzyme, indicating that tryptophan is important for enzyme activity. Residue 178 is the only conserved tryptophan near the substrate binding site, and thus it was a logical choice for mutation experiments. The location of tryptophan 178 on the three-dimensional crystal structure is shown in Figure 4. It is located near the edge of the binding pocket on the surface of the protein and is the first residue of ß-strand $ß_2S_1$ (nomenclature of Varghese et al., 1983). Since the orientation of the substrate bound in the

active site pocket is not known, it is not possible to predict what role this or any other residue may play in the catalytic reaction. Tryptophan 178 could have a role in either or both binding and hydrolysis of the substrate, but since binding is presumably a prerequisite for hydrolysis, the effect of the mutation of 178 Trp to Leu on the catalytic reaction could not be determined in the total absence of enzyme activity.

Loss of enzyme activity in the mutant could be accounted for by a change in the overall three-dimensional structure of the enzyme active site caused by this amino acid substitution. We do not believe this is the case for several reasons. It has been reported that proper transport and assembly of membrane proteins is at least partially dependent on the correct overall structure of the protein (Gething, 1985). The mutant protein is transported to the cell surface in a manner indistinguishable from wild-type. In immunofluorescence experiments using pooled monoclonal antibodies which are specific for the native NA structure, the mutant protein appeared identical to wild-type. In addition to experiments using chemical modification of active site residues (Polgar and Bender, 1966) several groups have done elegant site-specific mutagenesis experiments on enzymes with very encouraging results. Craik et al (1985) have mutated the active site of the trypsin molecule resulting in proteins with enzyme activity but altered substrate specificity. Similar experiments on tyrosyl-tRNA synthetase (Fersht et al., 1985) have elucidated the role of complementary hydrogen bonding in the binding of substrate. The X-ray crystal structure of mutant of influenza HA containing a single amino acid substitution has been determined and shows no structural changes from the wild-type except for the side chain substitution of glycine to aspartic acid (Knossow et al., 1984).

The SV40 expression system and site-specific mutagenesis procedure are powerful tools allowing us to study virtually any desired substitution. A single amino acid position can be changed to a number of different residues allowing us to study the characteristics of a particular side chain that are important for enzyme activity. If we can more fully understand NA activity at the molecular level, it may be possible to use the geometry of the active site to design specific inhibitors of NA which compete for the natural substrate.

Since the enzyme activity is common to all strains of the virus, such an approach may allow control of the disease across the wide variety of antigenically different strains.

ACKNOWLEDGMENTS

We are greatly indebted to Dr. Clayton Naeve and Dr. Robert Webster, St. Jude Childrens Research Hospital, Memphis, for synthesizing oligonucleotides and supplying antisera, respectively, and to Dr. Eric Hunter, University of Alabama at Birmingham for supplying vector DNA. This work was supported by Grant AI-19084 from NIAID. M. R. L. is supported by NIH Training Grant AI-07150. We thank Lynn Ritchie and Rodney Tucker for expert technical assistance and Eric Hunter and Peter Colman for many helpful discussions and for unpublished results.

Figure 1. Construction of the expression vector. Vector pQP was supplied by Dr. Eric Hunter. After deleting the Sal I site of pXf3 (closed triangle), a partial Cla I digest was performed and the linker 3'CAGCTGTAGC / CGGTCGACAT 3' was ligated into the open site. This created a new Sal I site and regenerated the Cla I site, resulting in vector pQPS. A wild-type or mutant cDNA clone of the NA gene of A/Tokyo/3/67 (N2) virus was inserted into the Sal I site. The plasmid sequences were removed by Bam HI digestion, the vector was recircularized, and the recombinant vector SVNA was transfected into CV-1 cells along with helper virus dl 1055 DNA. Abbreviations: PL, SV40 late promoter; PE, SV40 early promoter; ori, SV40 origin of replication; C, Cla I; S, Sal I; K, Kpn I. The open triangle represents the deleted Agno gene ATG initiator codon.

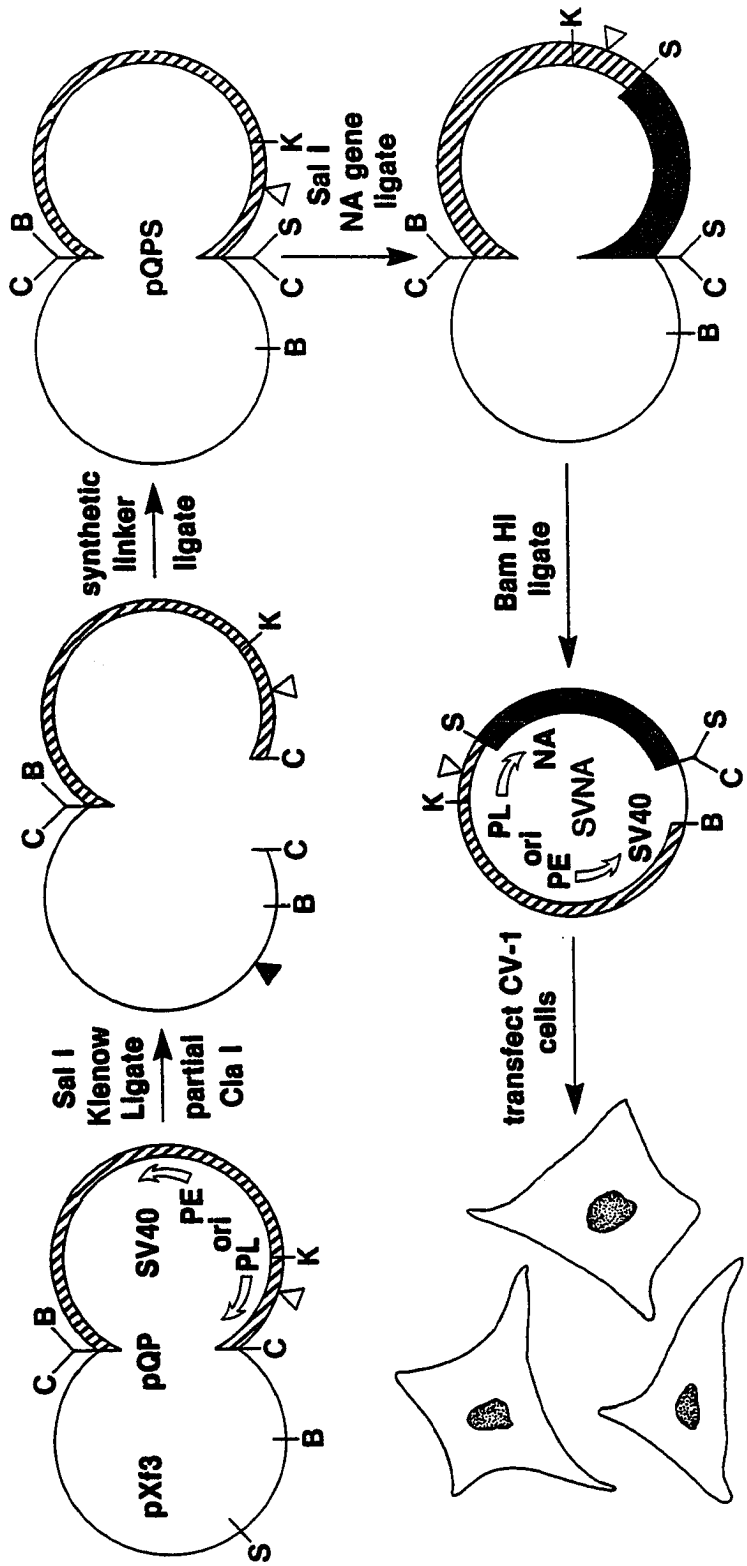


Figure 2. Indirect immunofluorescence for the detection of NA polypeptides. Intracellular fluorescence: (A) uninfected cells; (C) VH infected cells; (E) SVNAwt infected cells; (G) SVNA178L infected cells. Surface fluorescence: (B) uninfected cells; (D) VH infected cells; (F) SVNAwt infected cells; (H) SVNA178L infected cells.

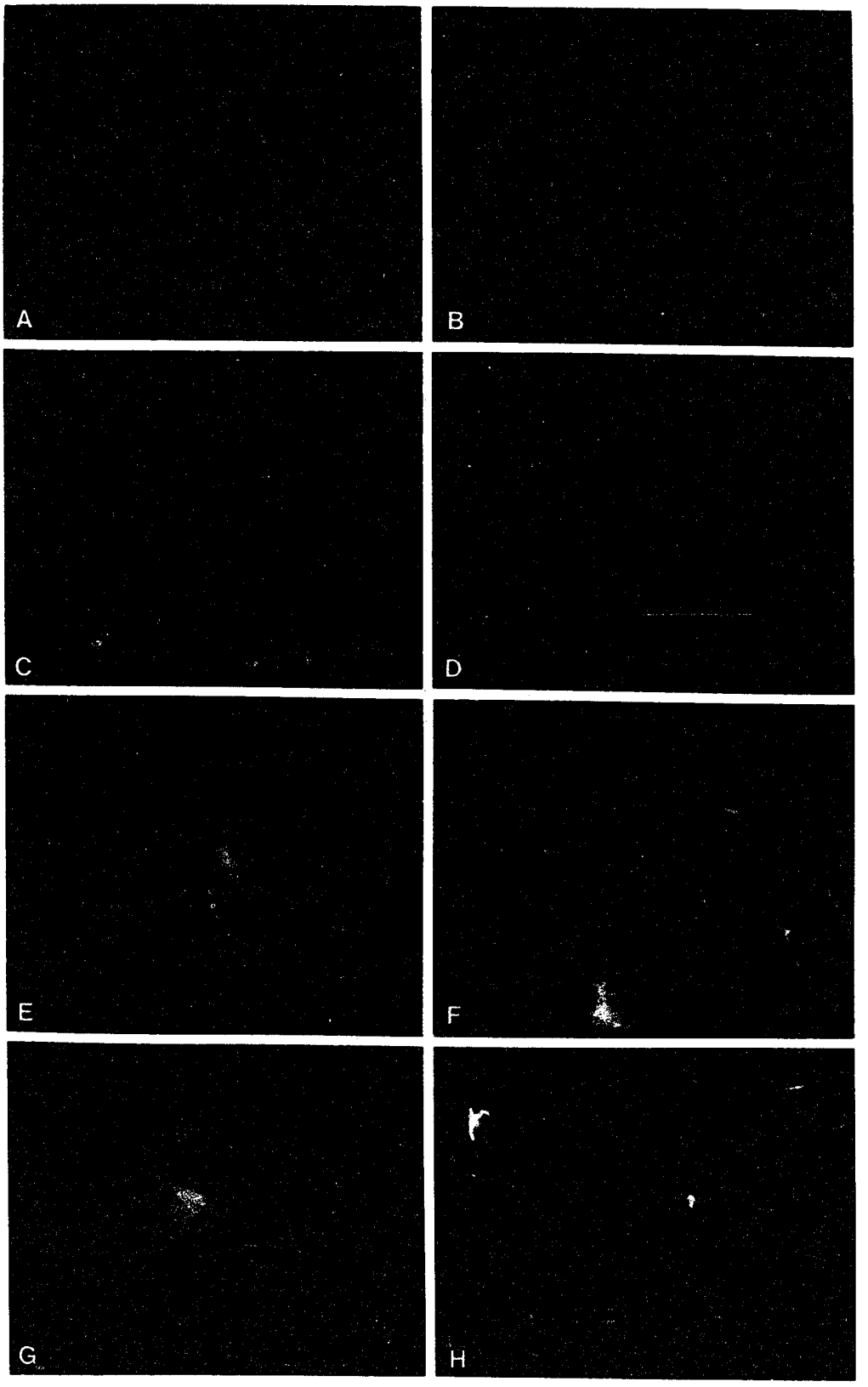


Figure 3. Assay for NA enzyme activity. Absorbance at 549 nm corresponds to the level of enzyme activity. Fetuin, a glycoprotein, and N-acetylneuraminyl lactose (NANL), a trisaccharide, were used as substrates. Results shown are the average of at least three different experiments.

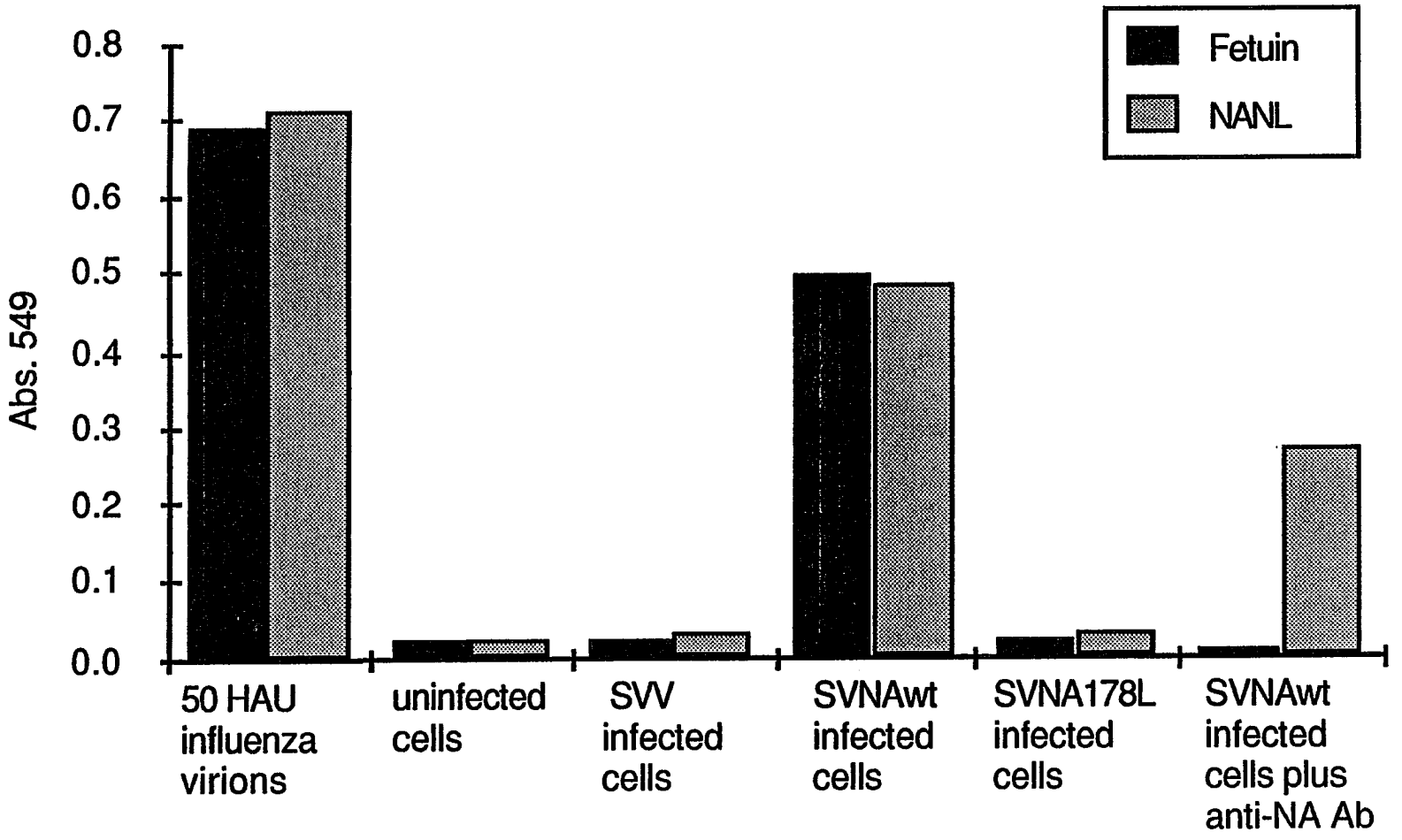
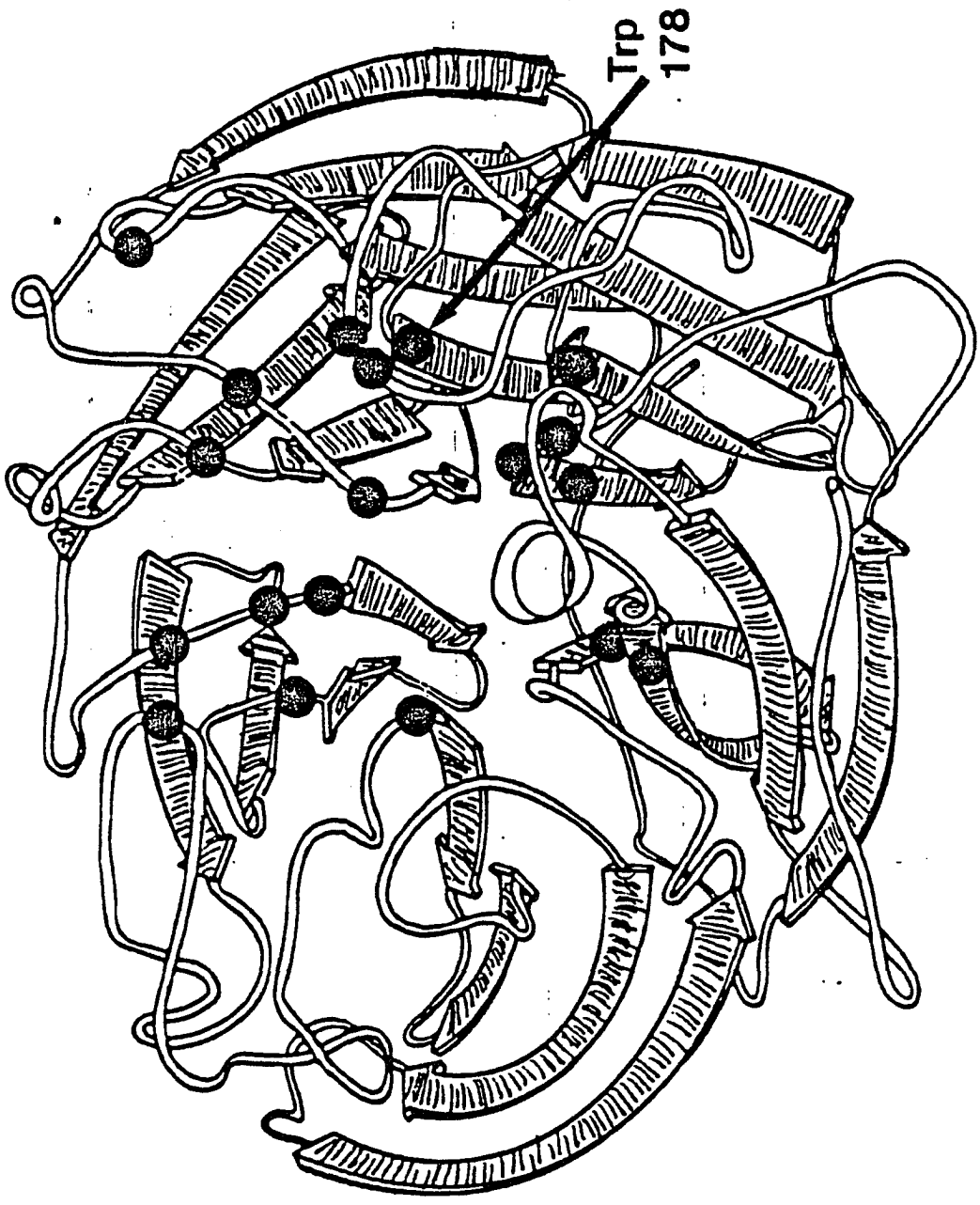


Figure 4. Position of the mutation on the three-dimensional structure of a monomer of NA. Filled circles represent the positions of conserved residues of the substrate binding pocket. The location of tryptophan 178, mutated to leucine, is indicated.



## REFERENCES

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

Aymard-Henry, M., Coleman, M. T., Dowdle, W. R., Laver, W. G., Schild, G. C., and Webster, R. G. (1973). Influenza virus neuraminidase and neuraminidase inhibition test procedures. Bull. WHO 48, 199-202.

Bachmeyer, H. (1972). Effect of tryptophan modification on the activity of bacterial and viral neuraminidase. FEBS Lett. 23, 217-219.

Basak, S., Tomana, M., and Compans, R. W. (1985). Sialic acid is incorporated into influenza hemagglutinin glycoproteins in the absence of viral neuraminidase. Virus Res. 2, 61-68.

Biggin, M. D., Gibson, T. J., and Hong, G. F. (1983). Buffer gradient gels and $^{35}S$ label as an aid to rapid DNA sequence determination. Proc. Nat'l. Acad. Sci. USA 80, 3963-3965.

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

Brady, J., Radonovich, M., Vodkin, M., Natarajan, V., Thoren, M., Das, G., Janik, J., and Salzman, N. P. (1982). Site-specific base substitution and deletion mutations that enhance or supress transcription of the SV40 late mRNA. Cell 31, 625-633.

Colman, P. M., Varghese, J. N., and Laver, W. G. (1983). Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature (London) 303, 41-44.

Craik, C. S., Largman, C., Fletcher, T., Roczniak, S., Barr, P. J., Fletterick, R., and Rutter, W .J. (1985). Redesigning trypsin: alteration of substrate specificity. Science 228, 291-297.

Davis, A. R., Bos, T. J., and Nayak, D. P. (1983). Active influenza virus neuraminidase is expressed in monkey cells from cDNA cloned in simian virus 40 vectors. Proc. Nat'l. Acad. Sci. USA 80, 3976-3980.

Fersht, A. R., Shi, J-P, Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., and Winter, G. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. Nature (London) 314, 235-238.

Fields, S., Winter, G. and Brownlee, G. G. (1981). Structure of the neuraminidase in human influenza virus A/PR/8/34. Nature (London) 290, 213-217

Gething, M. J. (ed.) (1985). "Protein Transport and Secrtetion," Current Communications in Molecular Biology Ser. Cold Spring Harbor Laboratory, N. Y.

Gething, M. J., and Sambrook, J. (1981). Cell surface expression of influenza hemagglutinin from a cloned DNA copy of the RNA gene. Nature (London) 293, 620-625.

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

Hunter, E., Hill, E., Hardwick, M., Bhown, A., Schwartz, D. E., and Tizard, R. (1983). Complete sequence of the Rous Sarcoma Virus env gene: identification of structural and functional regions of its product. J. Virol. 46, 920-936.

Joyce, C. M., and Grindley, N. D. F. (1983). Construction of a plasmid that overproduces the large proteolytic fragment (Klenow fragment) of DNA polymerase I of Escherichia coli. Proc. Nat'l. Acad. Sci. USA 80, 1830-1834.

Knossow, M., Daniels, R. S., Douglas, A. R., Skehel, J. J., and Wiley, D. C. (1984). Three-dimensional structure of an antigenic mutant of the influenza virus hemagglutinin. Nature (London) 311, 678-680.

Kozak, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes and organelles. Microbiol. Rev. 47, 1-45.

Laver, W. G. (1978). Crystallization and peptide maps of neuraminidase "heads" from H2N2 and H3N2 influenza virus strains. Virology 86, 78-87.

Laver, W. G., Air, G. M., Webster, R. G., and Markoff, L. J. (1982). Amino acid sequence changes in antigenic variants of type A influenza virus N2 neuraminidase. Virology 122, 450-460.

Lentz, M. R., Air, G. M., Laver, W. G., and Webster, R. G. (1984). Sequence of the neuraminidase gene of influenza virus A/Tokyo/3/67 and previously uncharacterized monoclonal variants. Virology 135, 257-265.

Luthman, H., and Magnusson, G. (1983). High efficiency polyoma DNA transfection of chloroquine treated cells. Nucl. Acids Res. 11, 1295-1308.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). "Molecular Cloning: a laboratory manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Markoff, L., Linn, B. -C., Sveda, M. M., and Lai, C. -J. (1984). Glycosylation and surface expression of the influenza neuraminidase requires the N-terminal hydrophobic region. Mol. Cell. Biol. 4, 8-16.

Norris, K., Norris, F., Christiansen, L., and Fiil, N. (1983). Efficient site-directed mutagenesis by simultaneous use of two primers. Nucl. Acids Res. 11, 5103-5112.

Palese, P., Tobita, K, Ueda, M., and Compans, R. W. (1974). Characterization of temperature-sensitive influenza virus mutants defective in neuraminidase. Virology 61, 397-410.

Pipas, J. M., Adler, S. P., Peden, K. W. C., and Nathans, D. (1980). Deletion mutants of SV40 that affect the structure of the viral tumor antigens. Cold Spring Harbor Symp. Quant. Biol. 44, 285-291.

Pipas, J. M., Peden, K. W. C., and Nathans, D. (1983). Mutational analysis of simian virus 40 T antigen: Isolation and characterization of mutants with deletions of the T-antigen gene. Mol. Cell. Biol. 3, 203-213.

Polgar, L., and Bender, M. L. (1966). A new enzyme containing a synthetically-formed active site. Thiolsubtilisin. J. Amer. Chem. Soc. 88, 3153-3154.

Rafelson, M. E., Scheir, M., and Wison, V. M. (1963). Studies on the neuraminidase of influenza virus. Arch. Biochem. Biophys. 103, 424-430.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Nat'l. Acad. Sci. USA 74, 5463-5467.

Schulman, J. L., and Palese, P. (1977). Virulence factors of influenza A viruses. J. Virol. 24, 170-176.

Seto, J. T., and Rott, R. (1966). Functional significance of sialidase during influenza virus multiplication. Virology 30, 731-737.

Tait, R. C., Rodriquez, R. L., and West, R. W. (1980). The rapid purification of T4 DNA ligase from a λT4 lig lysogen. J. Biol. Chem. 252, 813-815.

Varghese, J. N., Laver, W. G., and Colman, P. M. (1983). Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. Nature (London) 303, 35-40.

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

Wills, J. W., Srinivas, R. V., and Hunter, E. (1984). Mutations of the Rous Sarcoma Virus env gene that affect the transport and subcellular location of the glycoprotein products. J. Cell Biol. 99, 2011-2023.

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

Zoller, M. J., and Smith, M. (1983). Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. In: "Methods in Enzymology" (R. Wu, L. Grossman, and K. Moldave, eds.) Vol. 100, pp. 468-500. Academic Press, New York.

# SITE-DIRECTED MUTATION OF THE ACTIVE SITE OF INFLUENZA NEURAMINIDASE AND IMPLICATIONS FOR THE CATALYTIC MECHANISM

Michael R. Lentz,[1] Gillian M. Air,[1] Robert G. Webster,[2] and Peter M. Colman[3]

[1]Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294; [2]St. Jude Children's Research Hospital, 332 North Lauderdale, P.O. Box 318, Memphis, Tennessee 38101; and [3]CSIRO, Division of Protein Chemistry, 343 Royal Parade, Parkville 3052, Victoria, Australia

## SUMMARY

Conserved residues in the substrate-binding pocket of the influenza virus neuraminidase protein are likely to be involved in substrate binding or enzyme catalysis. A full-length cDNA clone of the neuraminidase gene from influenza A/Tokyo/3/67 was subcloned into an M13 vector and twelve amino acid substitutions were made in selected conserved residues using the oligonucleotide mismatch technique. The mutant neuraminidase genes were expressed from an SV40 vector and the proteins were assayed to determine the effect of the mutation on enzyme activity. The mutant enzymes are synthesized and transported to the cell surface identically to wild-type, however, some of the mutant proteins have completely lost enzyme activity, some mutants are fully active, while others have intermediate activity. The properties of the mutant enzymes are discussed together with implications for the catalytic mechanism.

# INTRODUCTION

Influenza virus has two surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA). Both proteins function in the immune recognition of the virus by the host and both undergo antigenic variation allowing viruses to escape previously established immunity to earlier strains. HA binds to host cell receptors and initiates penetration by fusing viral and host cell membranes. NA is an enzyme which catalyzes the cleavage of the α-ketosidic linkage between a terminal sialic acid residue and the adjacent residue on carbohydrate chains (Gottschalk, 1957). Although there is much structural and functional information known about the NA, the role of this enzyme during the course of viral infection is still not clear. It may have a general role of facilitating the movement of the virus within the host, first by allowing penetration of the mucous layer above the target respiratory epithelial cells (Burnet et al., 1947; Burnet, 1948) and subsequent movement from the infected cell by cleavage of the HA receptor on the cell surface (Seto and Rott, 1966; Webster and Laver, 1967; Palese et al., 1974). NA may also prevent aggregation of progeny viruses (Seto and Rott, 1966) and could unmask the activational cleavage site of HA (Schulman and Palese, 1977) by removing sialic acid from the side chains of the viral glycoproteins (Basak et al., 1985).

The biologically active part of the NA protein is a square, box-like head held above the viral envelope by a long narrow stalk anchored to the membrane at its N-terminus (Blok et al., 1982; Fields et al., 1981). Heads can be released from the virions by digestion with Pronase and retain all the biological properties of the intact protein (Laver, 1978; Laver et al., 1982). The three-dimensional structure of the NA heads from A/Tokyo/3/67 has been deduced from X-ray crystallographic studies. The protein is composed of four identical subunits, each made up of six, four-stranded β-sheets (Varghese et al., 1983). The product of catalysis, sialic acid, is observed to bind to a large pocket on the protein surface (Colman et al., 1983). A large number of amino acids line

this pocket, 24 of which are highly conserved in different types and subtypes of NA despite up to 50-60% variation between any two amino acid sequences. These residues are conserved across N1, N2, N7, N8, and N9 subtypes of influenza A for which sequence data is available as well as in influenza B (Air et al., 1985; Dale et al, in press). The location and conservation of these residues suggests a possible role in the enzyme activity of the protein. Inactivation of the function of the surface proteins may be a suitable approach to the control of influenza and we have therefore undertaken a more detailed study of the active site of the NA protein.

We have used oligonucleotide-directed mutagenesis to make amino acid substitutions in ten of the conserved residues in a cDNA clone of the NA of the A/Tokyo/3/67 (N2) strain of influenza A. The mutants and wild-type NA proteins were expressed from an SV40 late replacement expression vector and assayed for enzyme activity. In this paper we present the results of these experiments and discuss their implications for the catalytic activity of the NA.

## MATERIALS AND METHODS

DNAs, Viruses, and Cells. The cloning of a full-length cDNA copy of the NA gene of influenza virus A/Tokyo/3/67 has been reported (Lentz et al., 1984) and was the source of NA cDNA. The construction of SV40 vector pQPS used in these experiments has been described (Lentz and Air, 1986).

All recombinant plasmids were propagated in Escherichia coli strain RR1. CV-1 African green monkey kidney cells, permissive for SV40 replication, were used for expression of recombinant SV40 viruses. Cells were grown in Dulbecco's modified Eagle's medium containing 50 U/ml penicillin, 50 μg/ml streptomycin and 10% fetal calf serum.

Manipulation of DNAs. All manipulations of DNA were performed according to standard protocols (Maniatis et al., 1982). Synthetic oligonucleotides were synthesized by Dr. Jeffrey Engler on an Applied Biosystems DNA synthesizer.

Transfection and Virus Stocks. Recombinant SV40 virus stocks were prepared by transfection as described (Lentz and Air, 1986). High titer virus stocks were used at a dilution which produced complete cell death five days postinfection, generally 1:10 or 1:20.

Mutagenesis. The full-length clone of Tokyo/67 NA was subcloned into the Sal I site of M13mp18. Recombinant phage and replicative form DNA were grown in E. coli JM105 cells. The mutagenesis procedure of Zoller and Smith (1983) was used with modifications as described (Lentz and Air, 1986). Recombinant phage were screened for the mutations by plaque lift hybridization (Zoller, in press). Briefly, after transformation and plating cells onto bacterial lawns, plaques were grown overnight, then the plates placed on ice at 4° for one hr. A dry nitrocellulose filter was asymmetrically marked and placed onto the top agar and allowed to become thoroughly moistened for three to five min. Filters were removed slowly to prevent sticking of the top agar to the filter. After air drying for 15 min, the filters were baked under vacuum at 80° for two hr, then probed as described for phage supernatant dot blots (Zoller and Smith, 1983).

Immunofluorescence of Infected Cells. Internal and surface indirect fluorescent staining of infected cells was carried out as described (Wills et al., 1984) except that 3.7% Formalin in phosphate-buffered saline (PBS) was used instead of ethanol/acetic acid to fix surface stained cells. Two different anti-NA Ab preparations were used, a polyclonal rabbit antiserum and a pool of monoclonal Ab that specifically recognize the properly folded three-dimensional structure of Tokyo/67 NA (Lentz and Webster, unpublished results).

Radiolabeling, Immunoprecipitation, and Polyacrylamide Gel Electrophoresis of Proteins. Cells were labeled for 6 hr with D-(2-$^{3}$H)-mannose or 3 hr with L-[4,5-$^{3}$H(N)]leucine 3 days postinfection as described (Wills et al., 1984). Radiolabeled cells were lysed and the NA protein was immunoprecipitated (Hunter et al., 1983) using an excess of anti-Tokyo/67 NA polyclonal antisera preadsorbed to fixed CV-1 cells. Polyacrylamide gel electrophoresis and fluorography were carried out as described (Lentz and Air, 1986).

Neuraminidase Enzyme Assay. The colorimetric assay of Aymard-Henry et al (1973) was used. Five days postinfection, when cell death was nearly complete, the remaining cells and cell debri from two 60 mm plates infected with each recombinant vector were pelleted at 30,000 rpm for 30 min in an SW41 rotor. The pellet was resuspended in 200 µl of 0.25 mM $CaCl_2$, 0.8 mM $MgCl_2$, 0.15 M NaCl (CMS) and 100 µl of the suspension was transferred to each of two tubes. Fifty microliters of 40 mg/ml fetuin in 0.2 M sodium phosphate or potassium phosphate pH 6.0 was added to one set of tubes and 50 µl of 2 mM N-acetylneuraminyl lactose in the same buffer was added to the other set. The tubes were incubated at 37° for 4 hr after which time the N-acetyl neuraminic acid released by the enzyme was chemically converted to a pink chromophore, extracted into 1.5 ml 95% butanol / 5% acetic acid (v/v) and its aborbance read at 549 nm. Whole influenza virus (5 µl $10^4$ HAU/ml) was used as a source of enzyme as a positive control.

For enzyme assays over a pH range, cells from eight 60 mm plates infected with each recombinant vector were pelleted in a SW28 rotor for 40 min at 27,000 rpm and resuspended in 900 µl CMS. After distributing 100 µl of the cell suspension into 9 tubes, 50 µl of 40 mg/ml fetuin in 0.2 M potassium phoshate at the desired pH was added to each tube and the assay carried out as described above.

## RESULTS

Mutations of the Neuraminidase Protein. A total of twelve mutations were made, ten of which are in the conserved residues of the active site. The specific mutations and their location in the three-dimensional structure (Varghese et al., 1983) are shown in Figure 1 and are described below:

146 Asn to Ser; This is one of two mutated residues which is not located in the active site pocket. Asn 146 is the site of attachment of one of four N-linked carbohydrate side chains on the head of the Tokyo/67 NA (Ward et al., 1982, 1983) and this mutation abolishes this attachment site. This mutation was made to test the possibility that the WSN (N1) strain of influenza was neurovirulent because the loss of this carbohydrate from the

NA altered its enzyme activity. All other strains contain this carbohydrate side chain (Colman and Ward, 1985).

152 Arg to Lys and Ile; This charged residue is located on the top edge of the binding pocket with its side chain pointing directly into the pocket. Because of its location it appears to be a good candidate to be involved in binding substrate. Two mutations were made at this position, a conservative change from Arg to Lys which might provide information on the importance of the positive charge, and a change to Ile which is much less conservative, and alters both the charge and size of the side chain.

178 Trp to Leu; Bachmayer (1972) found that chemical modification of tryptophan in NA inactivated the enzyme which suggested that Trp might be involved in activity. Trp 178 is the only Trp in the vicinity of the substrate binding pocket and therefore is likely to be the only Trp involved in catalysis. A relatively conservative change to Leu was chosen to help maintain local structure possibly dependent on hydrophobic interactions. This mutant was used as template for a second mutagenesis to revert the Leu back to Trp as a control that the mutagenic procedures were not resulting in other mutations of the enzyme.

198 Asp to Asn; The N9 NA from influenza A/tern/Australia/G70C/75 (G70C) is able to hemagglutinate four times more efficiently than HA (Laver et al., 1984). This NA has Asn at residue 198 in place of the Asp in most strains (Air et al., 1985). To test the possibility that this Asn is responsible for the hemagglutinating and increased enzyme activity of this NA, the Asp in the Tokyo/67 strain was mutated to Asn as in G70C.

222 Ile to Val; This residue is well positioned to interact with the substrate and is one of the few hydrophobic residues that are conserved in the binding pocket. A conservative change was chosen to maintain the hydrophobic character of this side chain.

274 His to Tyr and Asn; His is an important residue in the active site of many enzymes, and residue 274 is the only His in the vicinity of the active site pocket. Since the pH optimum for NA (Mountford et al., 1982) and the $pK_a$ of His are very similar, it seemed possible that this His would be important for the activity of the NA.

277 Glu to Asp; This charged residue is located near the bottom of the binding pocket with its side chain pointing up into the pocket. A conservative change to Asp was made to preserve the negative charge at this position.

346 Thr to Asn; This mutation was designed to be a control since Thr 346 is not located in the active site pocket. The same substitution to Asn has been observed to occur in field isolates and therefore this mutation should not affect the enzyme activity.

371 Arg to Lys; This residue is located directly in the active site pocket with its side chain directed into the pocket (J. N. Varghese and P. M. Colman, personal communication). A conservative change of Arg to Lys was made to test the importance of this residue in the interaction with the substrate.

406 Tyr to Phe; This residue has recently been added to the list of conserved residues in the active site pocket (Air et al., 1985). A somewhat conservative change was made which maintains the aromatic nature of the side chain. The mutant residue is much less polar, however, and unable to participate in hydrogen bonding.

Expression of Wild-Type and Mutant Neuraminidase Proteins. NA proteins were expressed in CV-1 African green monkey kidney cells from an SV40 late replacement expression vector (Lentz and Air, 1986). Synthesis and transport of proteins was analyzed by immunoprecipitation of labeled cell lysates and by indirect immunofluorescence of infected cells.

Labeled NA proteins from infected cell lysates were precipitated using anti-NA rabbit antiserum, separated on 8% urea/phosphate/SDS gels and autoradiographed (Figure 2). In this gel system the reduced NA polypeptides migrate with an apparent molecular weight of 70,000 daltons and comigrate with NA from disrupted influenza virions. All of the mutant proteins have the same mobility as the wild-type with two exceptions. The protein with a mutation at residue 146 (Asn to Ser) which abolishes a carbohydrate attachment site migrates more quickly (mol wt 62,000) than the rest of the proteins in accord with the loss of a carbohydrate side chain. NA with mutant residue 406 (Tyr to Phe) appeared to migrate with an mol wt of 76,000. The reason for the apparent increase

in size of this protein is not known. The NA band from 198 Asp to Asn and the revertant consistently appeared slightly less intense than the rest of the NA proteins indicating less efficient expression from these two recombinants. SV40 VP1 protein was also less intense on the gel in these two vectors when L-[4,5-$^{3}$H(N)]leucine was used as a label (data not shown).

Indirect immunofluorescence was used to analyze the internal and surface expression of the NA proteins (Figure 3). Infected cells that were fixed and permeabilized prior to adding anti-NA rabbit antiserum stained brightly throughout the cytoplasm with increased intensity in the perinuclear region suggesting transport of the proteins through the Golgi apparatus. All of the mutants appeared identical to the wild-type NA. Uninfected cells or cells infected with a virus stock of recombinant vector and helper virus with no NA gene (VH) had only background fluorescence.

Surface fluorescence was done using two different anti-NA antibody preparations. Rabbit polyclonal antisera recognized both native and denatured NA proteins based on ELISA tests on whole native influenza virus and on boiled or detergent treated virus. A pool of monoclonal antibodies was only able to recognize NA in the properly folded three-dimensional structure using the same ELISA tests. Surface fluorescence with the rabbit antiserum revealed that the NA proteins were transported to the cell surface. There was no apparent difference in the fluorescent staining between wild-type and any of the mutants (Figure 3), although mutants 146 Asn to Ser and 406 Tyr to Phe had a lower percentage of surface fluorescing cells. These same amino acid substitutions, 146 Asn to Ser (loss of carbohydrate) and 406 Tyr to Phe, apparently affected the overall three-dimensional structure of the proteins as they were not recognized by the monoclonal Ab specific for the properly folded protein. The rest of the mutants appeared identical to wild-type with the monoclonal Ab (Figure 3).

Enzyme Activity of Wild-Type and Mutant Neuraminidase Proteins. The enzyme assays were performed on SVNA infected cells five days post-infection when cell death was nearly complete. This helped eliminate slight differences in the concentration of virus

stocks. Two substrates were used; N-acetylneuraminyl lactose, a trisaccharide, and fetuin, a highly sialated glycoprotein. Fetuin is more likely to mimic the natural substrate for the enzyme, whereas the trisaccharide is expected to have easier access to the active site. Assays were performed at pH 6.0, the optimum for the wild-type enzyme (Mountford et al., 1982). The standard colorimetric assay was used (Aymard-Henry et al., 1973) and the results are shown in Figure 4. Seven of the mutant proteins, with mutations 146 Asn to Ser, 152 Arg to Lys and Ile, 178 Trp to Leu, 198 Asp to Asn, 277 Glu to Asp and 406 Tyr to Phe have completely lost enzyme activity for both substrates. The rest of the mutants have activity that ranges from about four percent of wild-type (371 Arg to Lys) to fully active (222 Ile to Val and 346 Thr to Asn). Both mutations at residue 274 (His to Tyr and Asn) have about half wild-type activity. There were slight differences with the two substrates for some of the mutations, but no conclusions about the role of a particular mutated side chain could be deduced from these differences. When the inactive mutant at residue 178 was reverted back to the wild-type sequence, enzyme activity was regained, although activity was not quite as high as wild-type; this may result from the slightly lower expression described above.

pH Optimum of Wild-Type Neuraminidase and Two Mutants. Histidine 274 is the only histidine in the vicinity of the active site pocket. Since the pH optimum of the wild-type enzyme is near the $pK_a$ of histidine, and two mutations at this position (His to Tyr and His to Asn) both reduced the activity of the enzyme by about half, we decided to examine these mutations more closely. The enzyme assay was repeated on these mutants and the wild-type protein at pH values from 4.0 to 8.0 using fetuin as substrate. The results are graphed in Figure 5. The wild-type enzyme activity rose rapidly from pH 4.5 to a maximum of 5.5-6.0 then gradually fell off to 60-70% of maximum at pH 8.0. Both mutants had their pH optimum shifted to the acidic side, with maximum activity at pH 5.0. This activity was 52% and 54% of wild-type for the Tyr and Asn substitutions, respectively. Above the optimum pH, the activity of both mutant enzymes fell off rapidly to nearly zero (less than 10% of the maximum) by pH 8.0.

# DISCUSSION

Oligonucleotide-directed site-specific mutagenesis has been used to study specific details of a variety of enzymes. The substrate specificity of trypsin has been altered by this techique (Craik et al., 1985) and the role of complementary hydrogen bonding was analyzed in tyrosyl-tRNA synthetase (Fersht et al., 1985). The specific role of particular amino acids in the catalysis by dihydrofolate reductase (Howell et al., 1986) and carboxypeptidase A (Gardell et al., 1985) have been analyzed by creating specific mutations in these residues. We have used this techique to study the enzyme active site of the influenza virus neuraminidase protein. Little information was previously available about the active site other than its basic structure and which residues are conserved among the known sequences, therefore, we chose to make mutations in a large number of residues and use the information from all these mutants to shed light on the mechanism of catalysis or substrate binding.

Twelve mutations were made, ten of which are scattered around the pocket known to bind the substrate. The mutant NA genes were then cloned into an SV40 late replacement expression vector and NA proteins were expressed in CV-1 cells. The NA proteins anchored to the cellular membranes were assayed directly for their NA activity on two substrates, fetuin, a glycoprotein and N–acetylneuraminyl lactose, a trisaccharide. A wide range of activity was observed for the different mutants.

Two mutations apparently affected the three-dimensional structure of the protein as the mutant NA proteins were no longer recognized by a pool of monoclonal Ab specific for the properly folded protein. Both of these mutants are enzymatically inactive. Asn 146 mutated to Ser abolishes a carbohydrate attachment site which is present in all NA sequences except the neurovirulent WSN/33. The protein does migrate to the cell surface, but with reduced efficiency. It appears that this carbohydrate site is important for proper folding of the molecule in the N2 subtype.

The mutation at residue 406 from Tyr to Phe is also not recognized by the structure dependent monoclonal Ab. This protein is transported to the cell surface with reduced efficiency. The decreased polarity of the side chain at this position in the mutant may interfere with important interactions stabilizing the structure of the protein.

Mutants with substitutions at residues 152, 178, 198 and 277 have also completely lost activity, however the mutations do not appear to affect the transport or overall structure of these mutant proteins. Residue 152 was mutated from Arg to both Lys and Ile. Neither mutant had any activity indicating that the function of the Arg in the wild-type is very specific. It should be noted that while Arg to Lys is a conservative change, the Arg and Lys side chains are unequal in length.

Trp 178 mutated to Leu also loses all activity. Trp and Leu are both hydrophobic, however, the indole ring of Trp is both aromatic and able to participate in hydrogen bonding. Chemical modification of Trp in the NA also renders the enzyme inactive (Bachmeyer, 1972). It may participate in hydrogen-bonding or act as a donor in a charge-transfer interaction with the substrate. As a control, this mutant was reverted from Leu at residue 178 back to the wild-type Trp and enzyme activity was regained, indicating that no additional mutations had occured during the mutagenesis procedure.

Asp 198 was mutated to Asn and the mutant NA is inactive. Asn is found at the corresponding position in the N9 NA from G70C and it was therefore somewhat surprising that this mutant was totally inactive. The three-dimensional structure of N2 NA shows no obvious interactions involving Asp 198. The N9 NA with Asn at this position is able to agglutinate red blood cells, however, the mutant N2 NA with Asn is not. A reciprocal experiment in which the N9 NA of G70C was mutated at this residue to the Asp present in the Tokyo/67 wild-type enzyme was also performed. This mutant NA also completely lost enzyme activity and is unable to agglutinate red cells (Ritchie, Lentz, and Air, unpublished results), supporting the idea that the local environment surrounding these residues in the N2 and N9 neuraminidases is different, even though at the present resolution of the crystal structure we do not understand why.

The other substitution that inactivates the enzyme is 277 Glu to Asp. Both side chains have a negative charge at pH 6, but the side chain in the mutant is shorter than in the wild-type. This charged residue may participate in ionic interactions with the substrate in the wild-type, but be unable to make the necessary contacts for this interaction in the mutant. The Glu side chain might also be involved in stabilizing positively charged reaction intermediates.

Residue 371 was mutated from Arg to Lys and the mutant NA protein was found to have very low but detectable levels of activity, consistently from three to eight percent of the wild-type enzyme. The precise role of this side chain cannot be predicted without structural details of substrate bound in the active site.

Two mutants have activity that is identical to wild-type. At residue 222, a conservative mutation of Ile to Val was made. Both Ile and Val have rather inert side chains, so if an interaction with the substrate occurs, it must be a hydrophobic interaction. This residue could also have a structural role in the protein or the conservation of Ile at this position could be coincidence.

Residue 346 was mutated from Thr to Asn with no loss in enzyme activity. This was expected since this residue is not found in the active site pocket. This same substitution was observed to occur in field isolates of the virus and since the field isolates are all enzymatically active, this mutation serves as an additional control.

His 274 was changed to both Tyr and Asn. Both of the mutants behave identically, having 40-50% of wild-type activity at pH 6.0. These mutants were further characterized by assaying their activity over a range of pH from 4.0 to 8.0. The pH optimum for both mutants was shifted from 6 in the wild-type to 5. This shift in pH to the acidic side suggests a possible mechanism for the catalytic activity of this enzyme which has some elements in common with the mechanism of lysozyme, an enzyme which also catalyzes the hydrolysis of a glycosidic bond in polysaccharide substrates.

In lysozyme, Glu 35 and Asp 52 are both in close proximity to, but on opposite sides of the glycosidic bond cleaved in the reaction. Glu 35 is in a hydrophobic

environment and has an elevated $pK_a$ near 6, leaving this residue protonated at the pH at which the reaction occurs (pH 5) (Phillips, 1967; Blake et al., 1967). Glu 35 acts as a proton donor, passing its proton to the glycosidic oxygen and thereby breaking the bond. Asp 52, which is ionized at pH 5, stabilizes an intermediate carboxonium ion. A water molecule then protonates Glu 35 and provides a hydroxyl for the carboxonium ion to complete the reaction.

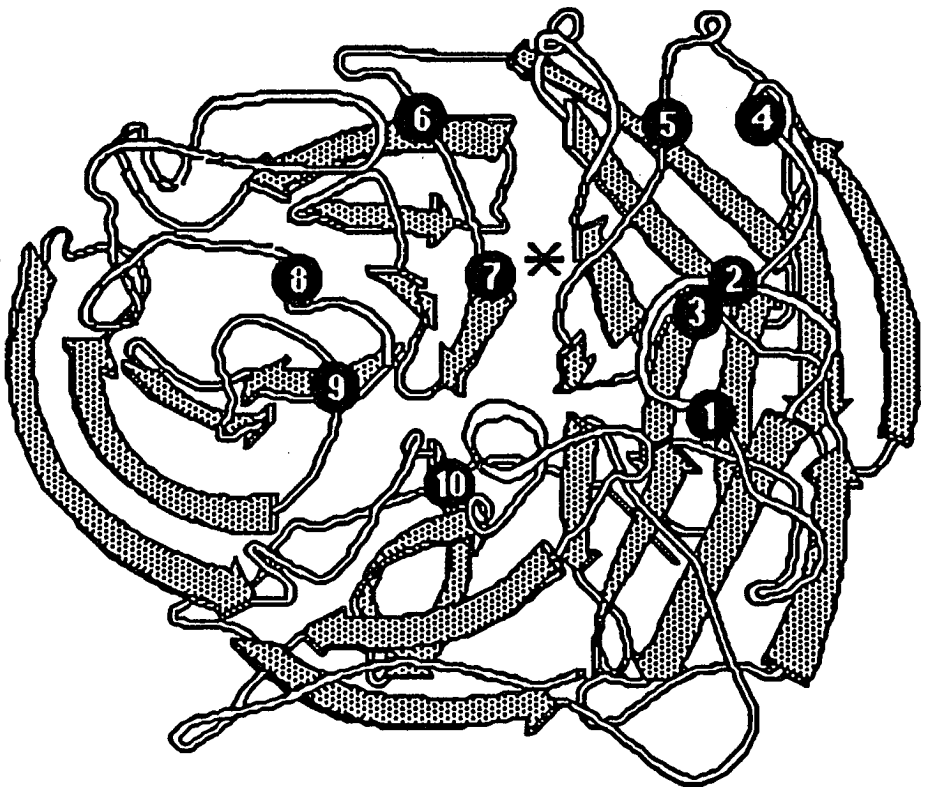
It was expected that His 274 may be involved in the catalytic reaction in NA, since the imidazole ring has a $pK_a$ very near the pH optimum for the reaction. It is unlikely that His 274 directly protonates the glycosidic oxygen, however, because enzyme activity is not abolished by substitution of His 274 to Tyr or Asn. It is more likely that His 274 elevates the $pK_a$ of nearby Glu 276 which may then be the proton donor for the reaction. In the absence of His at residue 274, the $pK_a$ of Glu 276 may resume a more typical value accounting for the lowered pH optimum of the reaction in the two mutants. Which of the many other acid groups in the active site is directly responsible for stabilizing the developing positive charge on a carboxonium ion intermediate is not clear, since it is not known precisely how the substrate is oriented in the active site during the reaction. Studies are in progress on the detailed interaction of sialic acid with the enzyme (J. N. Varghese and P. M. Colman, personal communication). Since sialic acid is only a weak inhibitor of the reaction (Colman et al., 1983), it is uncertain how clear the resolution of this structure interaction will be. It is likely that most conserved residues aside from the catalytically active His 274 and Glu 276 would be involved in binding the substrate in the active site, and studies on the interaction of sialic acid with the enzyme will help confirm this.

The enzyme activity of the neuraminidase protein is conserved in all the different strains and subtypes of the influenza virus. With detailed information about the activity and structure of the active site of this enzyme, it may be possible to design new inhibitors to effectively control the virus regardless of its antigenic structure.

ACKNOWLEDGMENTS

We are greatly indebted to Dr. Jeffrey Engler, University of Alabama at Birmingham, for synthesizing the mutagenic oligonucleotides. This work was supported by Grant AI-19084 from NIAID. M. R. L. is supported by NIH Training Grant AI-07150. We thank Lynn Ritchie and Dan Lakeman for expert technical assistance.

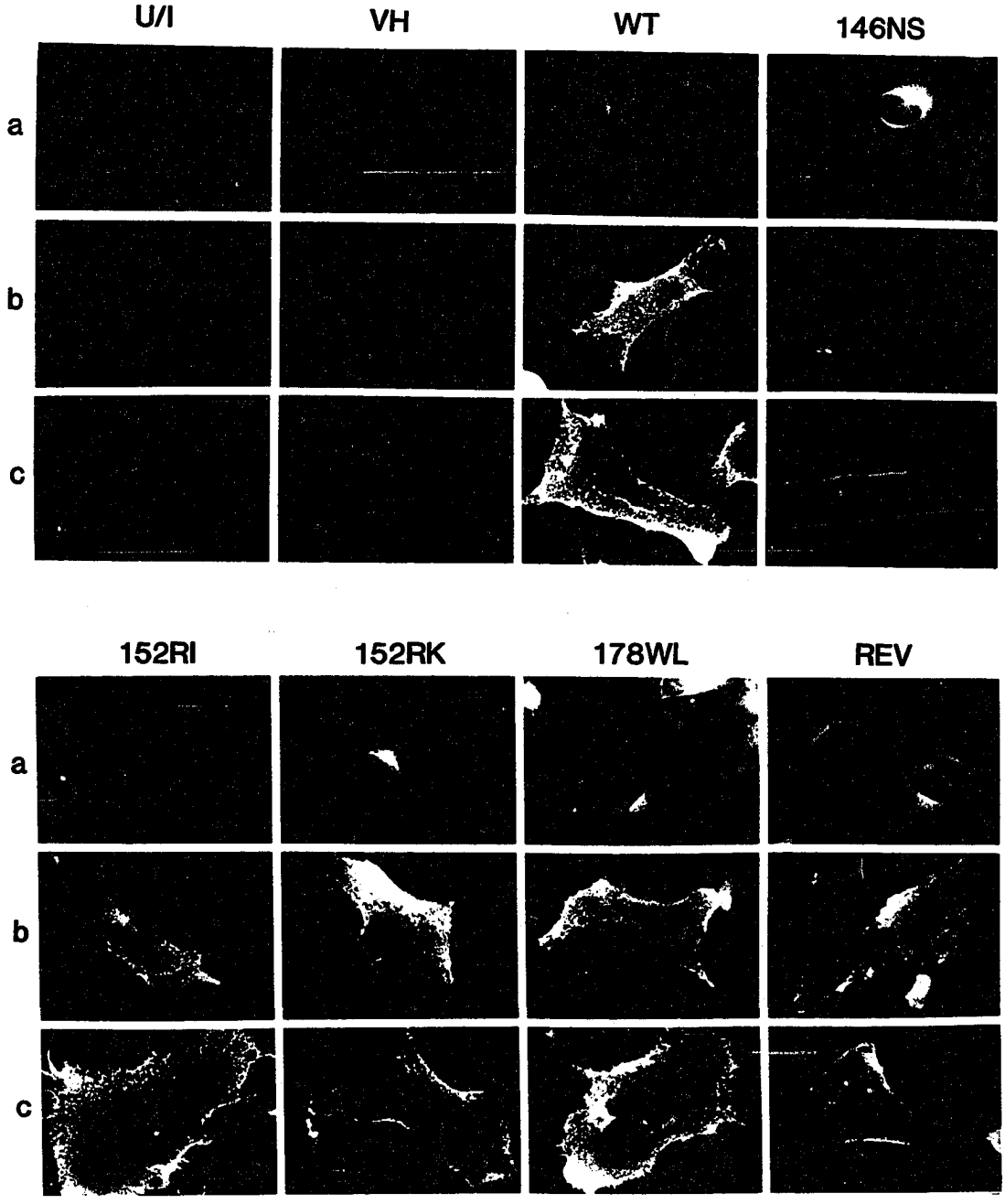
Figure 1. Location of mutations on the three-dimensional structure of NA. A schematic representation of a single monomer is shown. Amino acid substitutions are marked as closed numbered circles. A star represents the location of the sialic acid binding site.




1. 146 Asn to Ser
2.a. 152 Arg to Ile
2.b. 152 Arg to Lys
3. 178 Trp to Leu
4. 198 Asp to Asn
5. 222 Ile to Val
6.a. 274 His to Tyr
6.b. 274 His to Asn
7. 277 Glu to Asp
8. 346 Thr to Asn
9. 371 Arg to Lys
10. 406 Tyr to Phe

Figure 2. Immunoprecipitation of wild-type and mutant NA proteins. Infected cells were labeled with tritiated mannose as described in the text. Anti-Tokyo/67 NA rabbit antiserum was used to precipitate NA from infected cell lysates, then purified by binding to and eluting from Staphylococcus aureus protein A. Proteins were run on an 8% polyacrylamide gel, stained, and fluorographed. Whole disrupted $NWS_{HA}/Tokyo_{NA}$ influenza virions (FLU) were used as a marker for the NA protein. Lanes are marked STD for mol wt markers, U/I for uninfected control cells, VH for cells infected with the vector and helper virus stock with no cloned NA gene, and WT for cells expressing wild-type NA. REV is a revertant of the preceding mutant. Mutants are labeled with the amino acid position number followed by the single letter amino acid codes for the wild-type amino acid, then the amino acid substituted in the mutant.

Figure 3. Immunofluorescence of cells expressing NA proteins. Infected cells were treated with antibodies as described in the text. a, internal fluorescence of fixed, permeabilized cells using polyclonal rabbit antiserum as the primary antibody; b, surface fluorescence of unfixed cells using the same polyclonal antiserum as in a; c, surface fluorescence of unfixed cells using a pool of monoclonal antibodies which react exclusively with NA in the properly-folded three-dimensional structure. U/I, uninfected cells; VH, cells infected with the recombinant SV40 virus stock with no NA gene; WT, cells expressing wild-type NA; REV, cells expressing a revertant of the NA with the Trp to Leu mutation at residue 178. Mutants are labeled with the amino acid sequence number followed by the one-letter code for the wild-type amino acid, then the mutant amino acid.



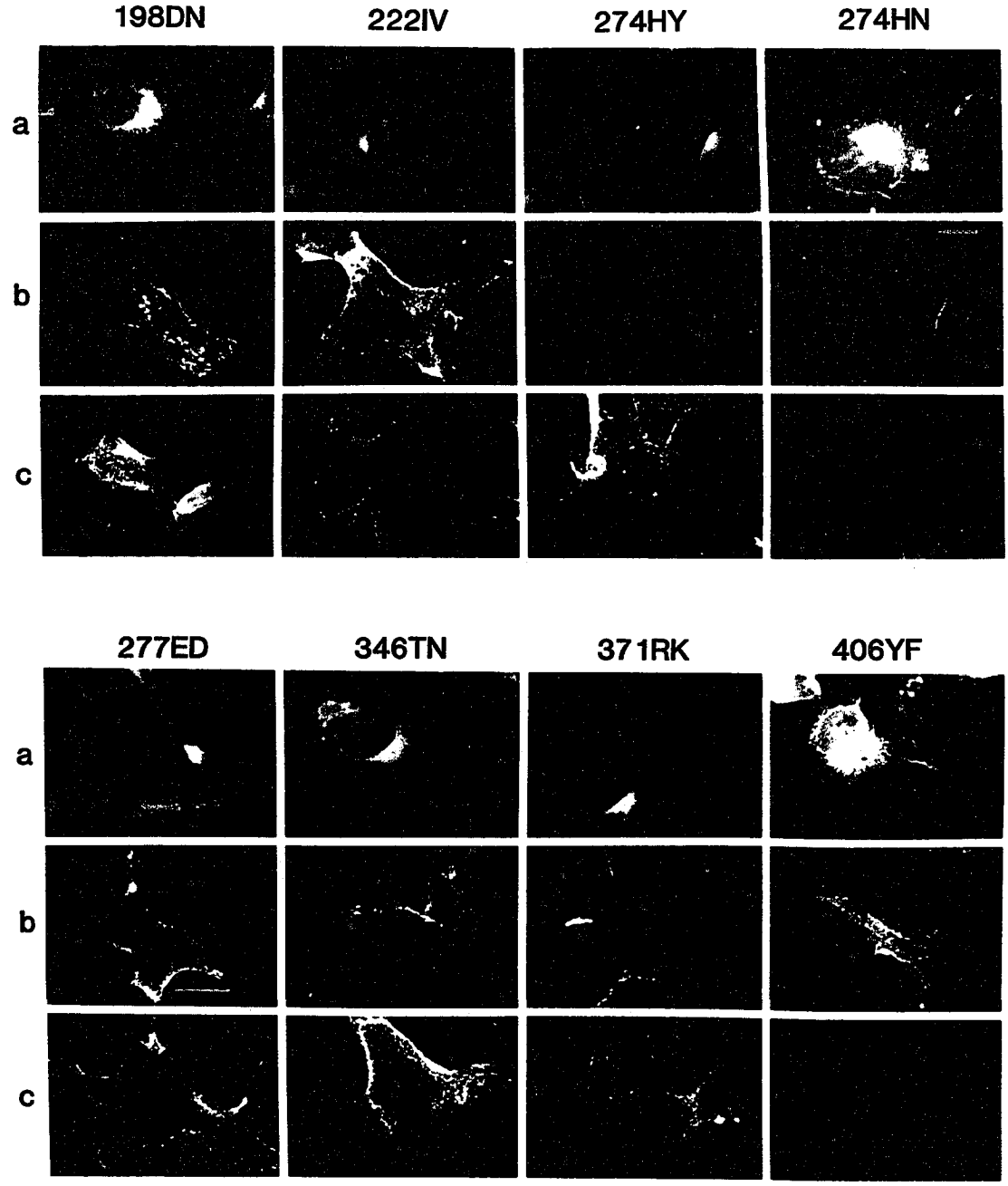
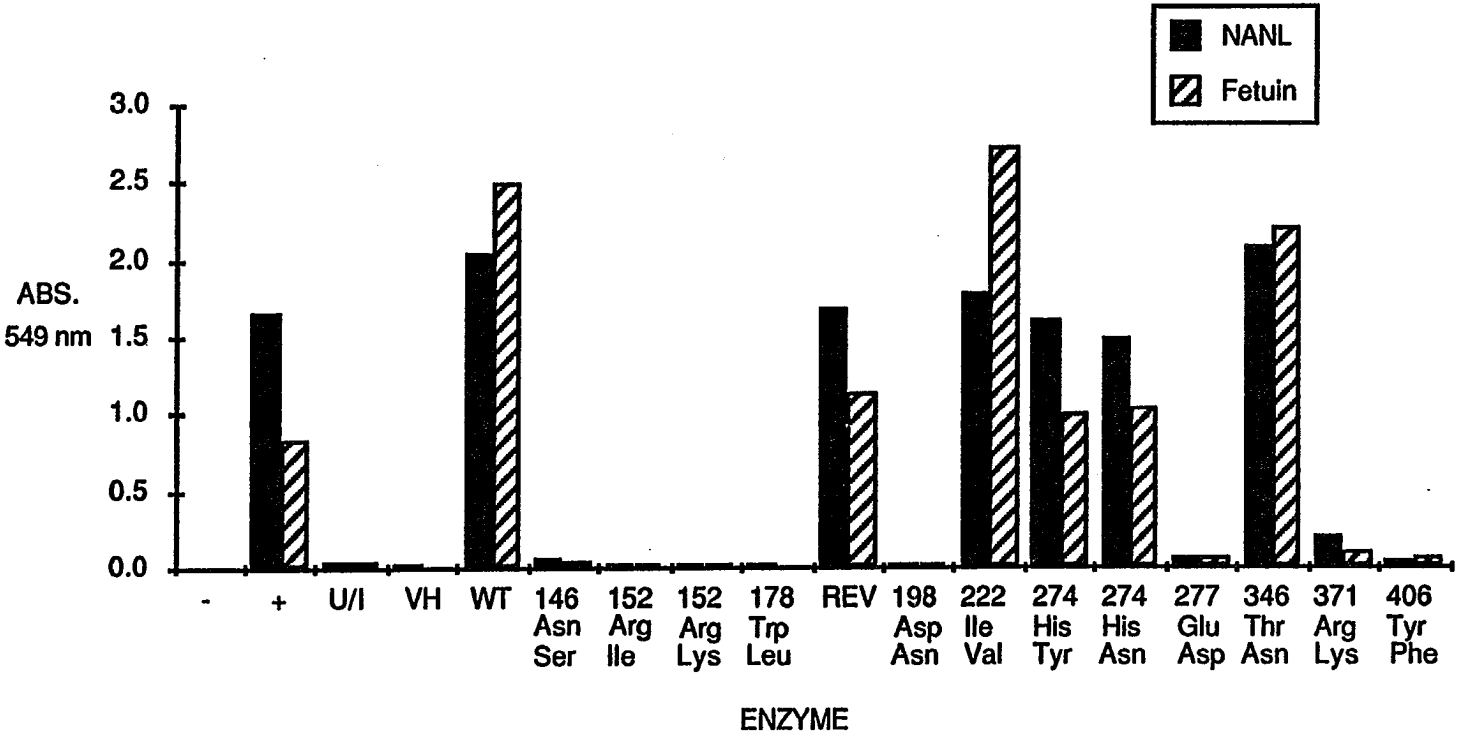


Figure 4. Enzyme activity of wild-type and mutant NA proteins. Enzyme activity was quantitated by determining the absorbance of the sample product at 549 nm. -, reagent blank; +, positive control using whole influenza virions as the source of NA; U/I, uninfected cells; VH, cells infected with recombinant SV40 virus stock with no NA gene; WT, cells expressing wild-type NA; REV, cells expressing a revertant of the 178 Trp to Leu mutation. Mutant proteins are labeled with the amino acid position above the wild-type amino acid. Below is the amino acid present at that position in the mutant. NANL; N-acetylneuraminyl lactose.



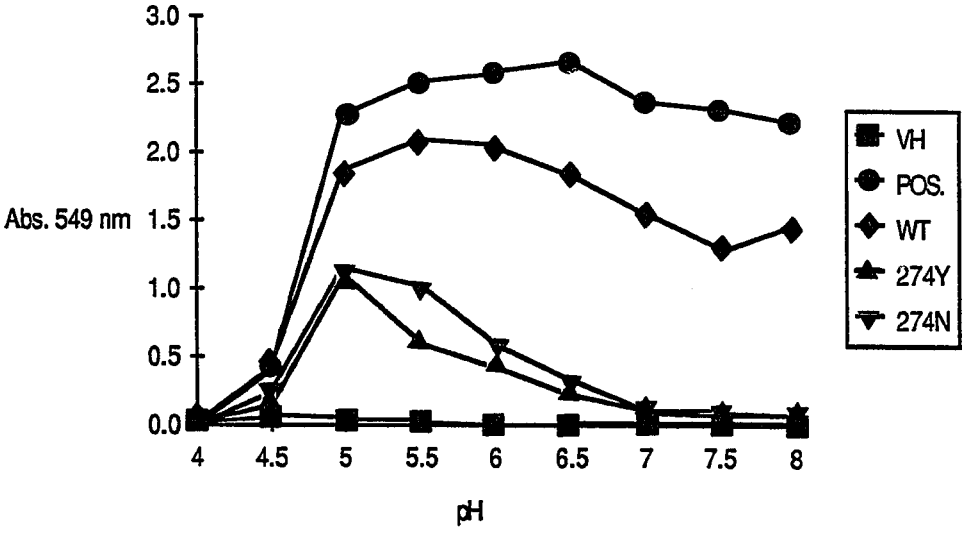
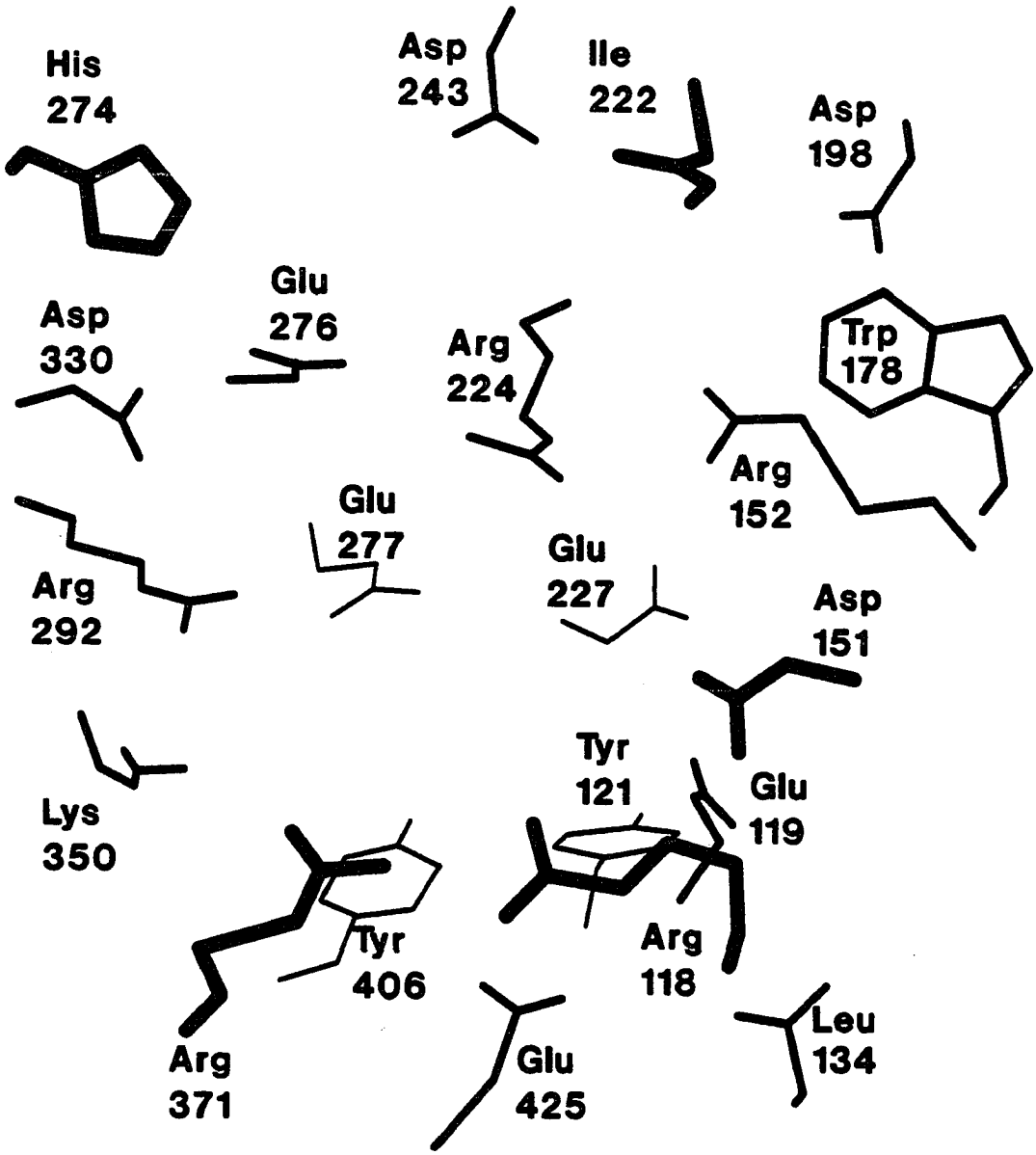



Figure 5. Enzyme activity of wild-type and His 274 mutants from pH 4.0 to 8.0. The absorbance at 549 nm corresponds to the level of enzyme activity. Pos, positive control using influenza virions as the source of NA; VH, recombinant SV40 vector with no NA gene; WT, cells expressing wild-type NA; 274HY, cells expressing NA with a substitution of Tyr for His at residue 274; 274HN, cells expressing NA with a substitution of Asn for His at residue 274.

Figure 6. Structure of the neuraminidase active site. The conserved residues located in the substrate binding pocket are shown. Mutations were made in residues 152, 178, 198, 222, 274, 277, 371, and 406. Side chains implicated in the catalytic mechanism are His 274 and Glu 276.



## REFERENCES

Air, G. M., Ritchie, L. R., Laver, W. G., and Colman, P. M. (1985). Gene and protein sequence of an influenza neuraminidase with hemagglutinin activity. Virology 145, 117-122.

Aymard-Henry, M., Coleman, M. T., Dowdle, W. R., Laver, W. G., Schild, G. C., and Webster, R. G. (1973). Influenza virus neuraminidase and neuraminidase inhibition test procedures. Bull. WHO 48, 199-202.

Bachmeyer, H. (1972). Effect of tryptophan modification on the activity of bacterial and viral neuraminidase. FEBS Lett. 23, 217-219.

Basak, S., Tomana, M., and Compans, R. W. (1985). Sialic acid is incorporated into influenza hemagglutinin glycoproteins in the absence of viral neuraminidase. Virus Res. 2, 61-68.

Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. (1967). Crystallographic studies of the activity of the hen egg-white lysozyme. Proc. Roy. Soc. 167, 378-388.

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

Burnet, F. M. (1948). Mucins and mucoids in relation to influenza virus action. IV. Inhibition by purified mucoid of infection and hemagglutinin with the virus strain WSE. Aust. J. Exp. Biol. Med. Sci. 26, 381-387.

Burnet, F. M., McCrea, J. F., amd Anderson, S. G. (1947). Mucin as substrate of enzyme action by viruses of the mumps influenza group. Nature (London) 160, 404-405.

Colman, P. M., Varghese, J. N., and Laver, W. G. (1983). Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature (London) 303, 41-44.

Colman, P. M., and Ward, C. W. (1985). Structure and diversity of the influenza virus neuraminidase. Curr. Topics Microbiol. and Immunol. 114, 177-255.

Craik, C. S., Largman, C., Fletcher, T., Roczniak, S., Barr, P. J., Fletterick, R., and Rutter, W. J. (1985). Redesigning trypsin: alteration of substrate specificity. Science 228, 291-297.

Dale, B., Brown, R., Miller, J., White, R. T., Air, G. M., and Cordell, B. (1986). Nucleotide and deduced amino acid sequence of the influenza neuraminidase genes of two equine serotypes. Virology, in press.

Fersht, A. R., Shi, J-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., and Winter, G. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. Nature (London) 314, 235-238.

Fields, S., Winter, G., and Brownlee, G. G. (1981). Structure of the neuraminidase in human influenza virus A/PR/8/34. Nature (London) 290, 213-217.

Gardell, S. J., Craik, C. S., Hilvert, D., Urdea, M. S., and Rutter, W. J. (1985). Site-directed mutagenesis shows that tyrosine 248 of carboxypeptidase A does not play a crucial role in catalysis. Nature (London) 317, 551-555.

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

Howell, E. E., Villafranca, J. E., Warren, M. S., Oatley, S. J., and Kraut, J. (1986). Functional role of aspartic acid-27 in dihydrofolate reductase revealed by mutagenesis. Science 231, 1123-1128.

Hunter, E., Hill, E., Hardwick, M., Bhown, A., Schwartz, D. E., and Tizard, R. (1983). Complete sequence of the Rous Sarcoma Virus env gene: identification of structural and functional regions of its product. J. Virol. 46, 920-936.

Laver, W. G. (1978). Crystallization and peptide maps of neuraminidase "heads" from H2N2 and H3N2 influenza virus strains. Virology 86, 78-87.

Laver, W. G., Air, G. M., Webster, R. G., and Markoff, L. J. (1982). Amino acid sequence changes in antigenic variants of type A influenza virus N2 neuraminidase. Virology 122, 450-460.

Laver, W. G., Colman, P. M., Webster, R. G., Hinshaw, V. S., and Air, G. M. (1984). Influenza virus neuraminidase with hemagglutinin activity. Virology 137, 314-323.

Lentz, M.R., and Air, G.M. (1986). Loss of enzyme activity in a site-directed mutant of influenza neuraminidase compared to expressed wild-type protein. Virology 148, 74-83.

Lentz, M. R., Air, G. M., Laver, W. G., and Webster, R. G. (1984). Sequence of the neuraminidase gene of influenza virus A/Tokyo/3/67 and previously uncharacterized monoclonal variants. Virology 135, 257-265.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Mountford, C. E., Grossman, G., Holmes, K. T., O'Sullivan, W. J., Hampson, A. W., Raison, R. L., and Webster, R. G. (1982). Effect of monoclonal antineuraminidase antibodies on the kinetic behavior of influenza virus neuraminidase. Mol. Immunol. 19, 811-816.

Palese, P., Tobita, K., Ueda, M., and Compans, R. W. (1974). Characterization of temperature-sensitive influenza virus mutants defective in neuraminidase. Virology 61, 397-410.

Phillips, D. C. (1967). The hen egg-white lysozyme molecule. Proc. Nat'l. Acad. Sci. USA 57, 484-495.

Schulman, J. L., and Palese, P. (1977). Virulence factors of influenza A viruses. J. Virol. 24, 170-176.

Seto, J. T., and Rott, R. (1966). Functional significance of sialidase during influenza virus multiplication. Virology 30, 731-737.

Varghese, J. N., Laver, W. G., and Colman, P. M. (1983). Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9Å resolution. Nature (London) 303, 35-40.

Ward, C. W., Elleman, T. C., and Azad, A. A. (1982). Amino acid sequence of the pronase-released heads of neuraminidase subtype N2 from the Asian strain A/Tokyo/3/67 of influenza virus. Biochem. J. 207, 91-95.

Ward, C. W., Murray, J. M., Roxburgh, C. M., and Jackson, D. C. (1983). Chemical and antigenic characterization of the carbohydrate side chains of an Asian (N2) influenza virus neuraminidase. Virology 126, 370-375.

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

Wills, J. W., Srinivas, R. V., and Hunter, E. (1984). Mutations of the Rous Sarcoma Virus env gene that affect the transport and subcellular location of the glycoprotein products. J. Cell Biol. 99, 2011-2023.

Zoller, M., and Smith, M. (1983). Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. In: "Methods in Enzymology". (R. Wu, L. Grossman and K. Moldave, eds). Vol 100, pp468-500. Academic Press, New York.

## SUMMARY AND DISCUSSION

Cloning of the A/Tokyo/3/67 Neuraminidase Gene. The preceding three papers describe the details of experiments designed to improve our understanding of the structure and function of the active site of the influenza virus neuraminidase protein. A full-length clone of the NA gene of the A/Tokyo/3/67 strain was prepared and sequenced as described in the first paper (Lentz et al., 1984). The Tokyo strain was chosen for a number of reasons. First, the sequence of this strain had not previously been determined at the nucleotide level and therefore the sequence would provide information useful in the study of antigenic shift and drift in the NA, especially since a number of monoclonal variants were prepared to this strain and their sequences determined (Webster et al., 1982; Laver et al., 1982; Lentz et al., 1984). More importantly, the Tokyo/67 strain was chosen for these studies since it was the strain used to determine the three-dimensional structure of the NA protein by X-ray crystallography. The 3Å structure of the Tokyo/67 strain was known, and since it forms high quality crystals, it was likely to be one of the more stable strains, a potential advantage when making mutations in the active site. Tokyo/67 NA is stable to Pronase digestion at 37° whereas some other N2 strains are inactivated by this treatment (Laver et al., 1982). N2 enzyme activity is also more resistant to heat than N1 strains (Drzeniek, 1972).

The three-dimensional structure of Tokyo/67 NA was recently solved (Colman et al., 1983; Varghese et al., 1983) and the positions of conserved residues were located on the structure making site-directed mutagenesis the most useful method of studying the active site. If the structure was not known, a better approach would have been to generate temperature-sensitive mutants of NA, then determine the mutations that led to this phenotype. This technique would have been more time consuming and given less precise information about the active site, since more than one substitution might occur in a given

mutant. Oligonucleotide-directed mutagenesis is a more powerful tool because a particular substitution can be made and this will be the only mutation present in the mutated gene.

Cloning the gene offered other advantages. Studying details of the NA active site would be easier without potential interference from other influenza gene products, so a heterologous expression system using a cloned gene was desirable. Genetic manipulations were made possible in some cases and much easier in others by using cDNA since double-stranded DNA is more ammendable to these manipulations than the single-stranded RNA found in the influenza virions. It is also easier to prepare large quantities of cloned DNA than viral RNA.

Expression of Neuraminidase from the Recombinant SV40 Virus Expression Vectors. Once the clone was prepared and sequenced, it was necessary to determine if enzymatically active NA protein could be synthesized from the gene. An SV40 recombinant vector system was chosen to direct the synthesis of NA from the cloned gene. This virus has a small, double-stranded DNA genome whose organization and control is well understood. It had already been used to successfully express a number of cloned genes, including influenza HA and NA (Gething and Sambrook, 1981; Davis et al., 1983; Markoff et al., 1984). Both the early and late promoters of SV40 drive the synthesis of large amounts of mRNA in infected cells. The virus productively infects monkey kidney cells and eventually results in cell death, so only transient expression is possible in this system. The release of progeny virions, however, means that recombinant virus stocks can be prepared and stored for repeated experiments. Other expression systems tend to be more variable; the SV40 system provides the consistent expression necessary for the study of wild-type and mutant NA proteins.

Three different vectors were constructed as described in the second paper (Lentz and Air, 1986). Two of these vectors expressed NA poorly or not at all. These vectors were not studied in detail, but some hypotheses about their inablility to express the cloned NA gene have been proposed. Vector pQDS was constructed by modifying pQD obtained from Dr. Eric Hunter (University of Alabama at Birmingham). In this vector, the late

coding region from the Hpa II to Bam HI site has been deleted and plasmid sequences were inserted at a Cla I site (formerly the Hpa II site). A Sal I site into which the NA gene could be cloned was inserted downstream of the bacterial sequences. When the bacterial sequences are removed with Cla I, the NA gene is positioned downstream of the late promoter. Only very weak expression of NA was observed from this construction. A clone of the matrix gene was also inserted into this vector with similar results. An analysis of the nucleotide sequence in the vicinity of the late promoter in this vector revealed two out-of-frame ATG codons between the late promoter (defined as the location of the unique Kpn I site of SV40) and the proper initiator codon of the NA gene. One of these ATGs is contributed by a Cla I linker sequence and the other is the initiator codon of the SV40 agno gene. The presence of these upstream AUGs may prevent efficient initiation of translation of NA protein by siphoning off ribosomes (Kozak, 1983).

The second vector was designed to eliminate these interfering codons. Vector pQKS was constructed by linearizing the plasmid vector at the unique Kpn I site, filling in the ends and adding a Sal I linker. The new Sal I site was used for inserting the NA gene. This vector was unable to express any detectable NA protein, probably because critical sequences of the late promoter were interrupted, preventing transcription of NA mRNA (Brady et al., 1982).

Vector pQPS was constructed to eliminate upstream AUGs and still have the cloning site downstream of the late promoter. pQP was supplied by Dr. Eric Hunter and was the starting material for this vector. pQP is similar to pQD except that the initiator ATG for the agno gene was eliminated (Perez et al., manuscript submitted). A custom linker containing a Sal I cloning site was inserted at the Cla I site in such a way that the Sal I site was upstream of the Cla I site and the bacterial sequences, so that the ATG of the Cla I linker would not interfere with expression of the NA gene. This construction is designed so that the bacterial sequences can be removed with either Cla I or Bam HI. Recent results have demonstrated that three Sal I linkers were inserted, the third in the opposite orientation, so that digestion with Sal I eliminates all of the Cla I sites. Therefore,

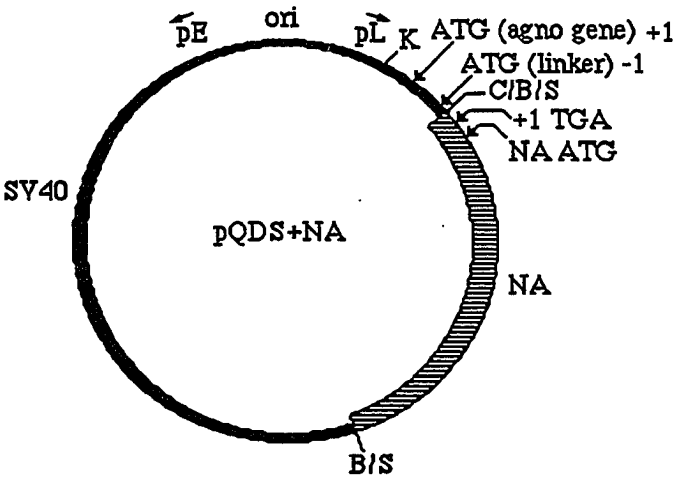
only Bam HI can currently be used to remove plasmid sequences (A-M. Hamilton, unpublished results). The three vector constructions are illustrated in Figure 1.

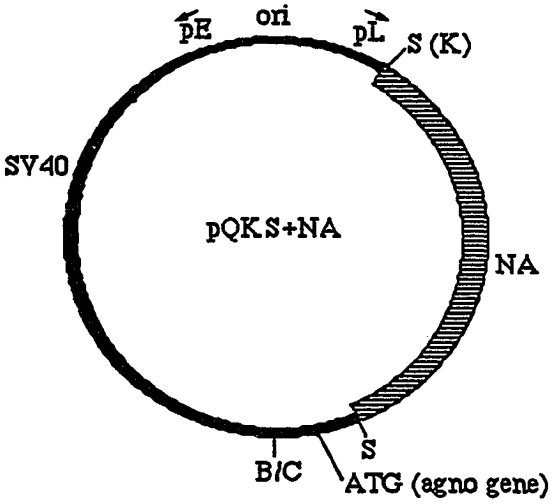
As described in the second paper, the pQPS vector construction results in high levels of expression of NA from the cloned gene (Lentz and Air, 1986). This was demonstrated by immuno-fluorescence of infected cells using anti-Tokyo/67 NA antibodies and by immunoprecipitation of radiolabeled proteins from infected cells using the same antiserum. The protein appears to pass through the Golgi apparatus where proteins are glycosylated; this was observed as bright fluorescence in the perinuclear region of infected cells. This same pattern of fluorescence was observed when N9 NA was expressed from the same vector, and the bright perinuclear stain corresponded to the region stained by wheat germ agglutinin (L. Ritchie and G. Air, unpublished results), known to specifically stain Golgi membranes (Virtanen et al., 1980). The NA protein is transported efficiently to the cell surface as demonstrated by immunofluorescence of live unfixed cells. This would be expected for the normal NA protein since influenza virus buds from the plasma membrane of the infected cell.

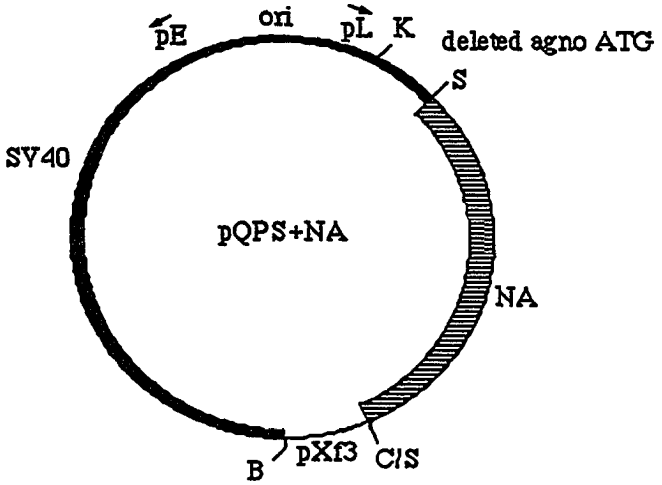
Infected cells expressing the wild-type NA gene have high levels of neuraminidase enzyme activity when either fetuin or N-acetylneuraminyl lactose are used as substrates. Uninfected cells or cells expressing only SV40 proteins are enzymatically inactive. The enzyme activity on cells expressing NA can be inhibited using anti-Tokyo/67 NA antibodies prior to the assay demonstrating that the activity on these cells is that of Tokyo/67 influenza virus NA and is a result of the expression of the cloned NA gene.

Other groups have expressed NA from similar vectors but only observed relatively weak levels of activity for low molecular weight substrates (Markoff et al., 1984; Davis et al., 1983). These differences in enzyme activity are difficult to explain. Different strains and subtypes of NA were used by each group and stability of the proteins may affect the activity seen from the expressed cloned gene. NA from the Tokyo/67 strain is believed to be relatively stable because of this proteins' ability to form good crystals. Heat stability of

Figure 1. SV40 vectors constructed for expression of neuraminidase. All vectors are diagrammed as DNA packaged into the recombinant SV40 virus particle. All vectors are modifications of plasmids constructed in the laboratory of Dr. Eric Hunter (University of Alabama at Birmingham). pQDS and pQKS were unable to express NA in detectable levels, whereas pQPS expressed NA well and was used in the expression studies of wild-type and mutant NA. Abbreviations: ori, SV40 origin of replication; pE, SV40 early promoter; pL, SV40 late promoter; NA, neuraminidase gene; B, Bam HI; C, Cla I; K, Kpn I; S, Sal I.



purified viral NA varies widely from strain to strain. N2 strains in general are the most heat stable while members of the N1 subtype are more heat labile (Drzeniek, 1972).

Other groups did not quantitate the amount of NA produced from their vectors, so a direct comparison of the expression levels is not possible. Low levels of expression from other vectors could account for reduced enzyme activity, however, all vectors are similar SV40 late replacement vectors and immunofluorescence and immunoprecipitation experiments indicate that all vectors express NA well. In one case (Davis et al., 1983), the wild-type gene was cloned in downstream of the agno gene ATG codon and an N-terminal deletion mutant was cloned in as a fusion peptide using the agno gene ATG as the start codon. Significantly higher levels of expression and enzyme activity were observed for the deletion/fusion NA than for the wild-type indicating that the agno gene ATG is a strong initiator compared to the downstream NA ATG in the wild-type construction. Markoff et al (1984) used a late replacement vector containing a 180 base pair deletion downstream of the late promoter, but the nature of this deletion was not described in detail. This data, together with the information derived from the vectors described above, indicate that ATGs between the late promoter and the desired start site significantly reduce the expression of cloned genes from SV40 late replacement vectors.

The expression of NA from pQPS was quantitated by comparing NA activity of infected cells to purified NA demonstrating that approximately $10^7$ enzymatically active NA molecules were present per infected cell. This is less than the $10^8$ HA molecules per infected cell using an SV40 vector reported by Gething and Sambrook (1981), where the total number of HA molecules per cell was determined by radioimmune assay on infected cell lysates. NA was quantitated by comparing the enzyme activity of a known number of infected cells to a known amount of purified Tokyo/67 NA heads as a standard, so only active NA molecules accessible to substrate were detected. The total number of NA molecules synthesized is likely to be somewhat higher than $10^7$ per cell.

Mutations in the Active Site of Neuraminidase. Since infected cells expressing wild-type NA had high levels of enzyme activity, the effect of mutations on the active site could be determined. Only in a few cases were there clues as to which residues lining the active site pocket might be interesting to mutate. There are 25 residues in this pocket which are conserved in all the sequenced strains. In the crystal structure of Tokyo/67 NA, all of the conserved residues appear to have their side chains oriented toward the pocket (Colman and Ward, 1985), but they are scattered all around the binding site. It is not known how the substrate is oriented when bound in the pocket for activity to occur. Trp 178 to Leu was the first mutation made. It was chosen because it is the only conserved Trp near the active site pocket and Trp has been shown to be important for activity by chemical modification studies (Bachmeyer, 1972). His is often found in enzyme active sites since it has a $pK_a$ near 6.0, in the pH range where most enzymes are active. Residue 274 is the only His in the binding pocket, so it was a logical choice for mutagenesis. The carbohydrate attachment site at 146 was chosen since this site is conserved in all known strains except WSN, a neurovirulent strain in which the neurovirulence was mapped to the NA. Residue 346 is neither conserved nor found in the active site pocket and therefore serves as a control. The substitution made at this position, Thr to Asn, is also observed to occur in field isolates. The rest of the mutations were chosen to account for residues in all parts of the pocket and to include a wide variety of side chain modifications. The details of the mutations are included in the third paper (Lentz et al., submitted).

The Active Site of Influenza Neuraminidase. A full range of enzyme activity was observed for all the mutants. A detailed study of the His 274 mutations has led to a model for the catalytic reaction that involves His 274, Glu 276 and possibly Glu 277. This would suggest that the primary role of all the remaining conserved residues is to bind and stabilize the substrate in the active site or to stabilize local structure in the enzyme molecule itself. It was observed that mutations of charged residues appeared to have a more dramatic affect on enzyme activity in general than polar or hydrophobic residues. Charge interactions may therefore play an important role in the binding and stability of the substrate. It is known

that the ionized carboxyl of sialic acid is critical for cleavage of the substrate, since any substitution at this position prevents cleavage (Drzeniek, 1973). Other positions in sialic acid are less critical; the N atom can be substituted with an acetyl group or a glycolyl group although activity towards the acetyl substitution is greater. Influenza NA is also active toward O-acetylated substrates (Drzeniek, 1973). The negative charge of the carboxyl may interact with one of the positive charged side chains conserved in the active site, an interaction which would be critical for activity. The precise enzyme-substrate interactions required for hydrolysis could best be determined by solving the crystal structure of substrate bound to the active site of the enzyme. This may not be possible since crystals of NA are still active and cannot be maintained at 4° for diffraction studies. The substrate would therefore be cleaved by the enzyme in the crystal, making the structure determination impossible. Sialic acid, the product of the reaction, does bind weakly to the enzyme and shows mild product inhibition. It can be soaked into crystals prior to diffraction analysis, but it is not known if sialic acid binds to the enzyme in precisely the same configuration as the substrate molecule.

One of the residues that was mutated, Ile 222, is conserved completely in all the known strains, yet the mutation had no apparent effect on the activity of the enzyme. The high degree of conservation when some strains show up to 50-60% total sequence variation suggests an important role for this side chain. The substitution, Ile to Val, is rather conservative, and perhaps a more dramatic change would have a more profound effect on enzyme activity. Alternatively, this side chain may be conserved merely to prevent steric or electrostatic hindrance of proper activity, although it might then be expected that this residue would demonstrate some degree of conservative variation in the field isolates. Additional mutations at this and other positions may provide information needed to sort out these possibilities. It is also possible that Val at position 222 might have a greater specificity for a particular glycosidic bond configuration, which would not be detected in these experiments since a mixture of $\alpha(2,3)$ and $\alpha(2,6)$ substrate molecules is present. Purified substrate molecules could be used to test this possibility. It is unknown how

many other conserved residues in the active site might play little or no role in the activity of the enzyme. Additional mutations of other residues in the active site would help answer some of these questions.

The Catalytic Mechanism of Neuraminidase. The proposed model for the mechanism of catalysis is based on the experiments described and the mechanism of lysozyme, which operates by general acid catalysis. Catalysis is most effective when the $pK_a$ of the active group is near the pH optimum of the enzyme. An acid of $pK_a$ 5 is a better catalyst than one of $pK_a$ 7, but at pH 7.0, 99% of a $pK_a$ 5 enzyme is ionized to the inactive form, while 50% of a $pK_a$ 7 acid is un-ionized and active. This accounts for the common involvement of His in such reactions, since it has a $pK_a$ of 6-7, near the optimum for many enzymes. His makes an important contribution to the activity of NA as well.

One crucial factor in determining the effectiveness of this type of catalysis is the ionization state of the enzyme under the reaction conditions. This is demonstrated very well in the case of lysozyme. The hydrolysis of substrate by lysozyme is dependent on two acid side chains, Glu 35 and Asp 52, amino acids which both have normal $pK_a$s of about 4.5. In the lysozyme molecule, however, these side chains are located in very different local environments, such that at pH 5, which is optimal for the enzyme, Asp 52 is ionized ($pK_a$ 4.5) while Glu 35 is not ($pK_a$ near 6). At pH 5, Glu 35 is an active proton donor/acceptor and contributes its proton to the C-O glycosidic bond in the substrate to initiate the cleavage.

General acid catalysis functions largely by stabilizing the transition state of the reaction. In the case of lysozyme, two critical elements come into play. The first is the presence of the ionized Asp 52 which stabilizes the charged carboxonium ion intermediate. The second factor is the distortion of the sugar residue in site D of the active site cleft, the residue on which the carboxonium ion forms. Proper binding of substrate into subsite D of the active site cleft appears to be dependent on the formation of the "half-chair" configuration of this sugar residue. This half-chair geometry promotes the formation of the carboxonium ion intermediate and may therefore play an important role in facilitating the

reaction. On the other hand, the carboxonium ion intermediate may form prior to complete binding, so that the enzyme is actually binding tightest to the reaction intermediate. In either case, in the process of binding, the enzyme forces the substrate to assume the geometry of the transition state (Phillips, 1967; Fersht, 1977; Stryer, 1981; Creighton, 1984).

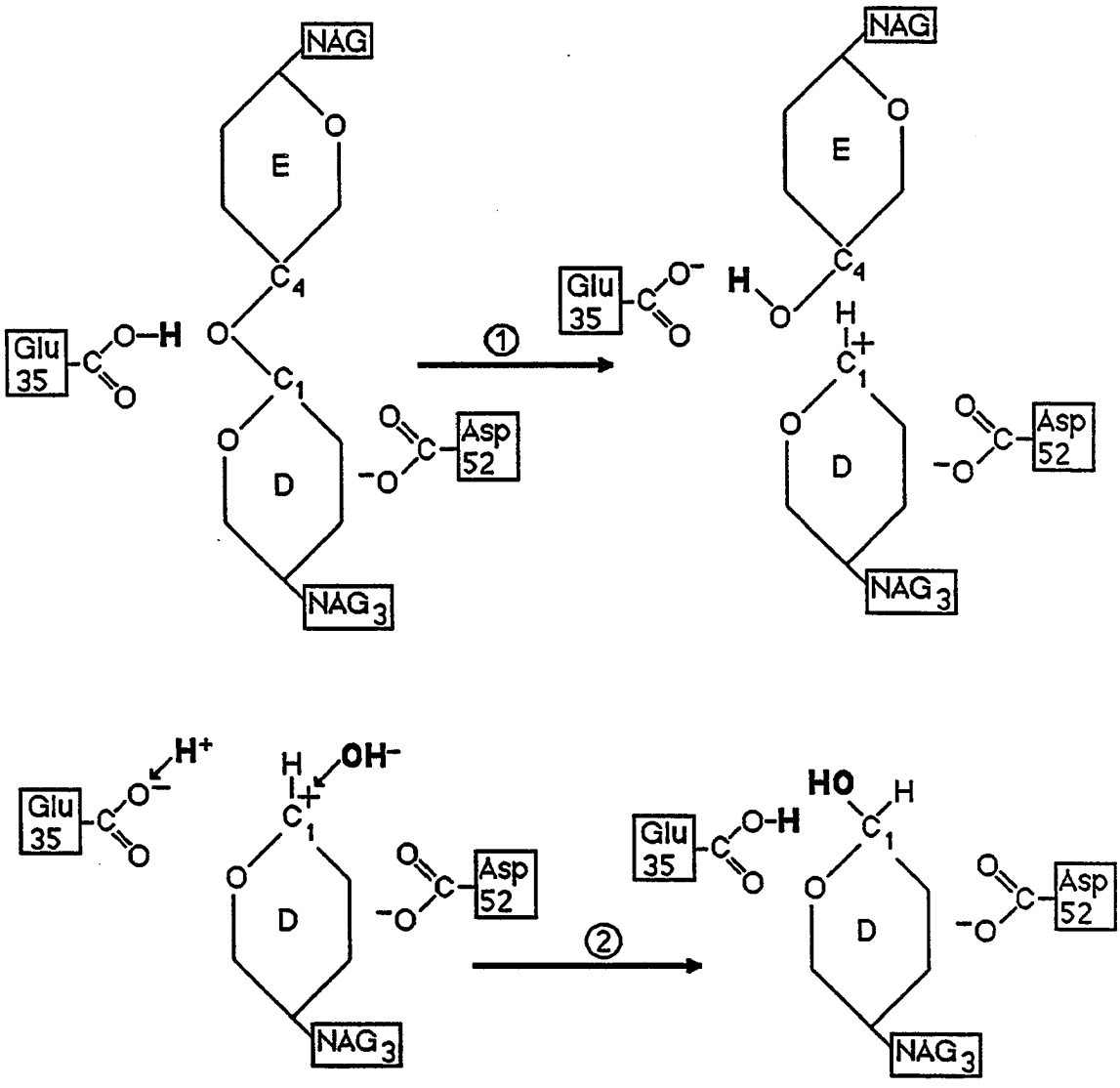
Once the substrate has bound to the active site cleft of lysozyme, the following steps occur:

1. A hydrogen ion is donated to the glycosidic C-O bond by the carboxyl group of Glu 35, thereby cleaving the bond. The disaccharide in subsites E and F can diffuse away. A carboxonium ion is formed by the positive charge on carbon 1 of the subsite D sugar. This is the transition state which is stabilized by the ionized carboxyl of Asp 52 and the tighter binding of the half-chair conformation of this sugar.

2. The hydrolysis reaction is completed when a water molecule contributes a hydroxyl to the carboxonium ion and a proton to the carboxyl of Glu 35. The tetrasaccharide from subsites A-D diffuses away. The enzyme is then ready for another round of catalysis. This mechanism is diagrammed in Figure 2.

This same general mechanism can be applied to the model proposed for catalysis by NA with a few modifications. In the NA molecule, Glu 276 functions to donate a proton to the glycosidic bond analogous to the Glu 35 of lysozyme. Glu 277 may provide a means of stabilizing an ionized intermediate as Asp 52 does in the lysozyme molecule. It is not known what ionization state either of these amino acids is in at the pH optimum of the reaction, but their location in the molecule appears to be exposed to the solvent (Colman et al., 1983; Varghese et al., 1983); this location and the requirement for His 274 at the normal pH optimum suggest that both of these side chains are ionized in the pH range of 6-7. In the case of the wild-type enzyme, the His residue provides a proton to the otherwise ionized Glu 276. Glu 276 is then able to donate this proton to the glycosidic bond resulting in cleavage. His 274 is not likely to donate the proton directly to the glycosidic bond since mutants of NA in which the His has been changed to Asn or Tyr are still active, although

Figure 2. The catalytic mechanism of lysozyme. A hexasaccharide is bound into subsites A through F of the enzyme, in this example, poly-N-acetylglucosamine (NAG). In step 1, the un-ionized carboxyl of Glu 35 donates a proton to the glycosidic bond, thereby breaking this bond. The disaccharide in subsites E and F can diffuse away. At the same time, a carboxonium ion intermediate forms at $C_1$ of the residue in subsite D, which is stabilized by the ionized carboxyl of Asp 52. In step 2, a water molecule contributes a proton to the carboxyl of Glu 35 and a hydroxyl to the carboxonium ion. After the product diffuses from the active site, the enzyme is ready for another round of catalysis. Adapted from Stryer, 1981.



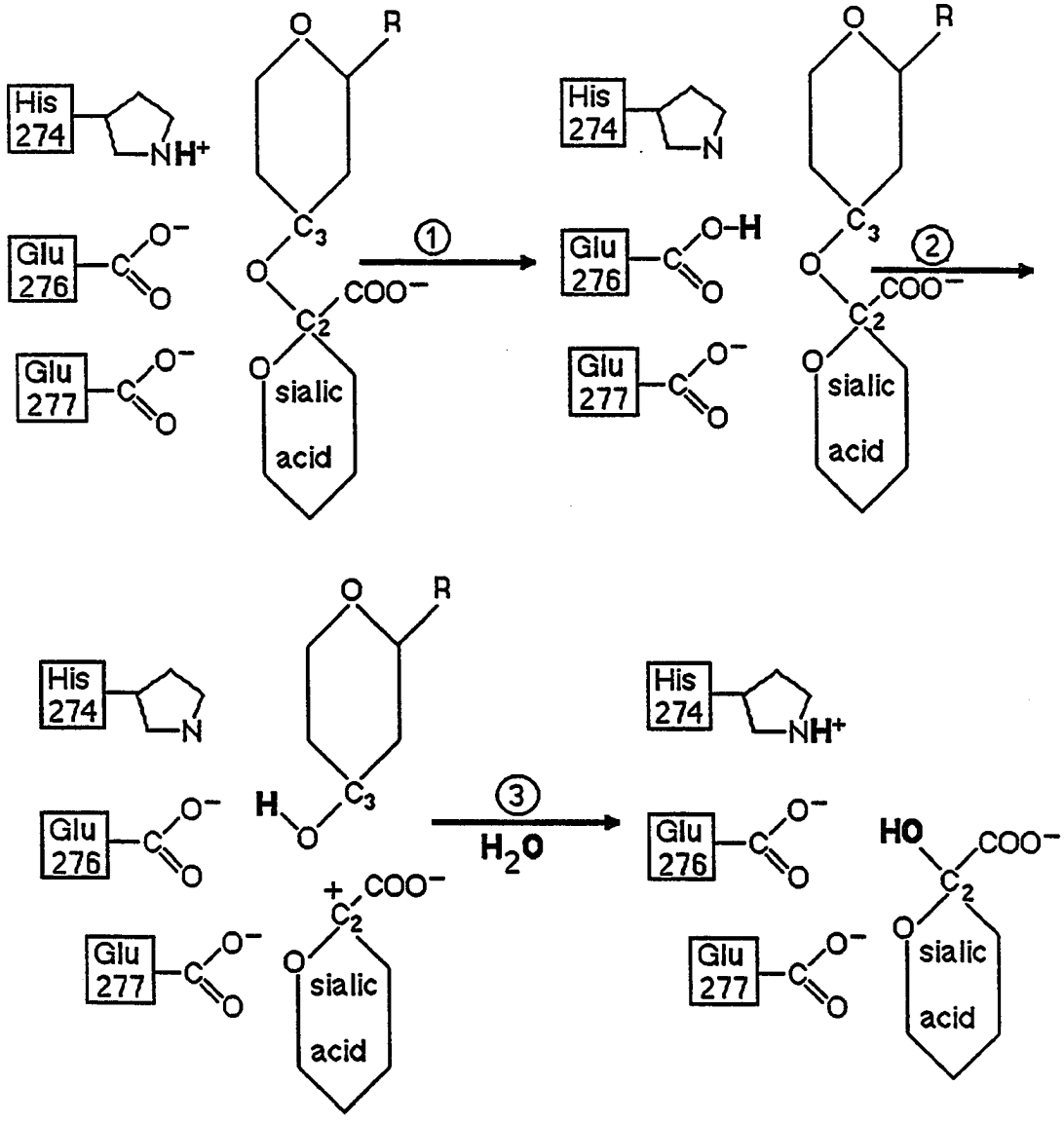
enzyme activity is reduced. After cleavage of the glycosidic bond, the polysaccharide chain minus the terminal sialic acid may diffuse away from the enzyme.

The pH optimum experiments on the His to Asn and His to Tyr mutants lend support to the roles of His 274 and Glu 276 in the proposed mechanism. The pH optimum is shifted to the acidic side (from pH 6 to pH 5) in both of these mutants. This would be expected if the His was no longer available to provide a proton, since the pH optimum is now closer to the $pK_a$ of Glu than in the wild-type enzyme. A lower pH is required for activity since Glu 276 would be a more effective proton donor at this lower pH in the absence of His 274.

The presence of a carbonium ion intermediate has not been demonstrated for NA, however for the proposed mechanism to operate, such an intermediate is likely. It could potentially be stabilized by the ionized carboxyl of Glu 277 which is positioned adjacent to Glu 276, presumably in the vicinity of the substrate cleavage site. Experiments described in the third paper showed that a conservative mutation of Glu to Asp at residue 277 abolishes enzyme activity, indicating that this residue has an important role in the catalytic mechanism. The shorter side chain in this mutant may be unable to make the electrostatic contribution necessary to stabilize the transition state. Any of the other conserved acidic residues of the active site could potentially help stabilize the charged intermediate. The weaker binding of sialic acid to the enzyme compared to substrate suggests that a half-chair conformation of the transition state may be forming as in lysozyme. This conformation could be the form which binds most tightly to the active site, and thereby allows for preferential binding of substrate and only weak product inhibition by sialic acid.

In the final step, a water molecule provides a hydroxyl to the carbonium ion intermediate and reprotonates the His residue, after which free sialic acid diffuses from the enzyme. The enzyme is then ready for another round of catalysis. This model for the catalytic mechanism is diagrammed in Figure 3.

Figure 3. The catalytic mechanism of neuraminidase. The substrate contains sialic acid linked α(2,3) to any polysaccharide chain (R). After binding to the active site, His 274 donates a proton to the ionized side chain of Glu 276 as shown in step 1. In step 2, Glu 276 uses this proton to break the glycosidic bond, resulting in the release of the polysaccharide minus its terminal sialic acid. A possible ionized intermediate forms, which may be stabilized by the ionized carboxyl of Glu 277 or another conserved acidic residue in the active site. In the final step, a water molecule reprotonates His 274 and contributes a hydroxyl to the transition state intermediate, resulting in free sialic acid which can diffuse from the active site. The enzyme is then ready for another round of catalysis.



Future Directions. The experiments described in this dissertation have provided a molecular view of the active site of influenza NA. The data presented along with the mechanism of lysozyme have provided a model for the catalytic mechanism for the NA molecule. This is only a model, however, and new experiments must be designed and performed in order to prove or disprove this model.

The proposed mechanism has provided new direction as to which amino acid residues are most likely to contribute useful information if mutated. Clearly Glu 276, the proton donor in the model, should be mutated to several different residues. If it plays the important role suggested by the mechanism, any non-conservative mutation should abolish activity. More conservative mutations such as Asp or Gln may modify activity and provide more detailed information on the role of Glu in the wild-type enzyme. It might be possible to shift the pH optimum to the basic side by mutating residue 276 to Arg or Lys, although the larger side chain may sterically interfere with the proper binding of the substrate. A variety of conserved residues distant from the catalytic center may have critical functions in the binding of substrates. Conservative and non-conservative mutations in a number of these residues along with computer modeling may be necessary to determine the precise configuration of the enzyme-substrate complex.

Another angle from which to approach the testing of the mechanism is to use different mutants of the enzyme with a variety of substrates which contain subtle modifications. Some NA mutants may have enhanced activity to a particular glycosidic bond configuration or have greater or reduced specificity for a particular oxygen or nitrogen substitution. It may be difficult to find or prepare some useful modified substrates.

Crystallographic studies of the enzyme complexed to sialic acid or its derivatives could be very useful in determining the structure of the enzyme-substrate complex. The model suggests that a carbonium ion intermediate may be the conformation bound most efficiently by the enzyme; sialic acid may therefore not be the best compound to use in these crystals. The importance of the half-chair configuration for the lysozyme intermediate was

supported by binding studies using a lactone analog of the substrate that has the same geometry as the transition state before it is bound to the enzyme (Stryer, 1981). Binding and crystallographic studies with such an analog of sialic acid could be very useful in determining the geometry of the enzyme-substrate complex.

The role of particular side chains in the reaction could be tested by determining the ability of mutants to bind inhibitor or substrate molecules. If a particular amino acid was important for binding of the substrate to the active site, a mutation at this position should reduce the ability of the protein to bind inhibitors or substrates. If mutations are made in residues that are part of the catalytic mechanism, the ability of the protein to bind the substrate is less likely to be affected. These experiments were attempted using tritium-labeled 2'-deoxy-2,3-dehydro-N-acetylneuraminic acid, a known inhibitor of NA and tritium-labeled N-acetylneuraminyl lactose, a substrate for the enzyme. These compounds were incubated with cells expressing wild-type NA or cells infected with the SV40 vector without the NA gene. No conclusions could be drawn since the background of radioactivity on cells with no NA was always as high as cells expressing NA under a variety of incubation conditions. The process of radiolabeling appears to have altered the structure of these compounds since the inhibitor reacts like sialic acid in the enzyme assay, whereas unlabeled inhibitor does not. These experiments can be repeated using different compounds or similar ones labeled more carefully so as not to interfere with their structure. The conditions need to be set to reduce the background on mock infected cells, so that the data obtained from mutant NA proteins can be interpretted.

It is hoped that detailed information on the mechanism and geometry of the active site of influenza NA will provide a useful starting point for the rational design of potential inhibitors of this enzyme. The enzyme and its activity is conserved in all the influenza viruses, even though antigenic properties of this protein are not. Such inhibitors might therefore provide a universal means of controlling the spread of this serious and sometimes deadly virus.

References

Bachmeyer, H. (1972). Effect of tryptophan modification on the activity of bacterial and viral neuraminidase. FEBS Lett. 23, 217-219.

Brady, J., Radonovich, M., Vodkin, M., Natarajan, V., Thoren, M., Das., G., Janik, J., and Salzman, N. P. (1982). Site-specific base substitution and deletion mutations that enhance or suppress transcription of the SV40 major late RNA. Cell 31, 625-633.

Colman, P. M., Varghese, J. N., and Laver, W. G. (1983). Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature (London) 303, 41-44.

Colman, P. M., and Ward, C. W. (1985). Structure and diversity of influenza virus neuraminidase. Curr. Topics Microbiol. Immunol. 114, 177-255.

Creighton, T. E. (1984). "Proteins: Structures and Molecular Properties." W. H. Freeman and Co., New York.

Davis, A. R., Bos, T. J., and Nayak, D. P. (1983). Active influenza virus neuraminidase is expressed in monkey cells from cDNA cloned in simian virus 40 vectors. Proc. Nat'l. Acad. Sci. USA 80, 3976-3980.

Drzeniek, R. (1972). Viral and bacterial neuraminidases. Curr. Topics Microbiol. Immunol. 59, 35-74.

Drzeniek, R. (1973). Substrate specificity of neuraminidases. Histochem. J. 5, 271-290.

Fersht, A. (1977). "Enzyme Structure and Mechanism." W. H. Freeman and Co., Reading/San Fransisco.

Gething, M-J., and Sambrook, J. (1981). Cell-surface expression of influenza hemagglutinin from a cloned DNA copy of the RNA gene. Nature (London) 293, 620-625.

Kozak, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes and organelles. Microbiol. Rev. 47, 1-45.

Laver, W. G., Air, G. M., Webster, R .G., and Markoff, L. J. (1982). Amino acid sequence changes in antigenic variants of type A influenza virus N2 neuraminidase. Virology 122, 450-460.

Lentz, M. R., and Air, G. M. (1986). Loss of enzyme activity in a site-directed mutant of influenza neuraminidase compared to expressed wild-type protein. Virology 148, 74-83.

Lentz, M. R., Air, G. M., Laver, W. G., and Webster, R. G. (1984). Sequence of the neuraminidase gene of influenza virus A/Tokyo/3/67 and previously uncharacterized monoclonal variants. Virology 135, 257-265.

Markoff, L., Linn, B. -C., Sveda, M. M., and Lai, C. -J. (1984). Glycosylation and surface expression of the influenza neuraminidase requires the N-terminal hydrophobic region. Mol. Cell. Biol. 4, 8-16.

Phillips, D. C. (1967). The hen egg-white lysozyme molecule. Proc. Nat'l. Acad. Sci. USA. 57, 484-495.

Stryer, L. (1981). "Biochemistry." Second Edition. W. H. Freeman and Co., San Fransisco.

Varghese, J. N., Laver, W. G., and Colman, P. M. (1983). Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9Å resolution. Nature (London) 303, 35-40.

Virtanen, I., Ekblom, P., and Laurila, P. (1980). Subcellular compartmentalization of saccharide moieties in cultured normal and malignant cells. J. Cell Biol. 85, 429-434.

Webster, R. G., Hinshaw, V. S., and Laver, W. G. (1982). Selection and analysis of antigenic variants of the neuraminidase of N2 influenza viruses with monoclonal antibodies. Virology 117, 93-104.

GRADUATE SCHOOL
UNIVERSITY OF ALABAMA AT BIRMINGHAM
DISSERTATION APPROVAL FORM

Name of Candidate Michael R. Lentz

Major Subject Molecular Cell Biology

Title of Dissertation Characterization of the Active Site of the Influenza Virus Neuraminidase Protein

Dissertation Committee:

, Chairman

Director of Graduate Program

Dean, UAB Graduate School

Date 8/22/86

PS-1428