Panyim, S., & Chalkley, R. (1969a) *Biochem. Biophys. Res. Commun. 37*, 1042.
Panyim, S., & Chalkley, R. (1969b) *Biochemistry 8*, 3972.
Pehrson, J. R., & Cole, R. D. (1981) *Biochemistry 20*, 2298.
Pieler, C., Adolf, G. R., & Swetly, P. (1981) *Eur. J. Biochem. 115*, 329.
Pragnell, I. B., Arndt-Jovin, D. J., Fagg, B., & Ostertag, W. (1980) *Exp. Cell Res. 125*, 459.
Rastl, E., & Swetly, P. (1978) *J. Biol. Chem. 253*, 4333.
Scher, W., & Friend, C. (1978) *Cancer Res. 38*, 841.
Smith, B. J., & Johns, E. W. (1980) *Nucleic Acids Res. 8*, 6069.
Smith, B. J., Walker, J. M., & Johns, E. W. (1980) *FEBS Lett. 112*, 42.
Tarnowka, M. A., Baglioni, C., & Basilico, C. (1978) *Cell (Cambridge, Mass.) 15*, 163.
Terada, M., Nudel, V., Fibach, E., Rifkind, R. A., & Marks, P. M. (1978) *Cancer Res. 38*, 835.
Tobey, R. A., & Ley, K. D. (1970) *J. Cell Biol. 46*, 151.
Tobey, R. A., & Ley, K. D. (1971) *Cancer Res. 31*, 46.
Tobey, R. A., & Crissman, H. A. (1972) *Exp. Cell Res. 75*, 460.
Tobey, R. A., Petersen, D. F., Anderson, E. C., & Puck, T. T. (1966) *Biophys. J. 6*, 567.
Varshavsky, A. J., Bakayev, V. V., & Giorgiev, G. P. (1976) *Nucleic Acids Res. 3*, 477.
Walters, R. A., Gurley, L. R., & Tobey, R. A. (1974) *Biophys. J. 14*, 99.
Walters, R. A., Tobey, R. A., & Hildebrand, C. E. (1976a) *Biochem. Biophys. Res. Commun. 69*, 212.
Walters, R. A., Tobey, R. A., & Hildebrand, C. E. (1976b) *Biochim. Biophys. Acta 477*, 36.
Whitlock, J. P., Jr., & Simpson, R. T. (1976) *Biochemistry 15*, 3307.
Wigler, M. H., & Axel, R. (1976) *Nucleic Acids Res. 3*, 1463.
Xue, S., & Rao, P. N. (1981) *J. Cell Sci. 51*, 163.
Zlatanova, J. S. (1980) *FEBS Lett. 112*, 199.
Zlatanova, J. S. (1981) *Mol. Cell. Biochem. 35*, 49.

Variation in the Membrane-Insertion and "Stalk" Sequences in Eight Subtypes of Influenza Type A Virus Neuraminidase†

Janet Blok*,‡ and Gillian M. Air§

ABSTRACT: The two membrane-bound surface antigens of influenza virus, hemagglutinin (HA) and neuraminidase (NA), are known to vary considerably in amino acid sequence. From immunological studies nine serologically distinct subtypes of neuraminidase have been characterized, yet these all exhibit the same enzyme activity. Many studies have been designed to investigate variation in the antigenic properties of the hemagglutinin and/or the neuraminidase. Here we have investigated the sequence variation of regions of the influenza neuraminidase which are not involved in antigenicity. These regions are the hydrophobic transmembrane segment and the "stalk". The NA gene of at least one strain from each of eight of the nine NA subtypes was sequenced from the 3′ end by using the dideoxy method in order to examine these regions. The results reveal that the predicted protein sequence at the N terminus is identical for the first six amino acids in all subtypes while the next six are the same in most subtypes. Following the first 12 amino acids, there is virtually no conservation of particular amino acid side chains in the transmembrane sequence and in the stalk region of the neuraminidase, although the character of the polypeptide is maintained. The transmembrane segment contains a high proportion of hydrophobic amino acids while the stalk region is rich in potential glycosylation sites, and it also contains at least one Cys residue which may form intermolecular disulfide bonds in the neuraminidase tetramer (at variable positions in the different subtypes). The different predicted protein sequences of the eight NA subtypes can therefore be accommodated into four regions of the tetrameric structure of the neuraminidase, which is compatible with the electron micrographs: an N-terminal conserved hexapeptide, the hydrophobic transmembrane segment, a thin stalk which is stabilized by carbohydrate and intermolecular disulfide bonds, and the enzymatically and antigenically active "head". Genetic variation of the influenza neuraminidase is not confined to antigenic properties.

The enzyme neuraminidase (NA)[1] (EC 3.2.1.18) is one of the two projecting surface antigens on particles of influenza viruses, the other being the hemagglutinin. Neuraminidase hydrolyzes terminal *N*-acetylneuraminic acid (NANA) from receptors for the virus located on the surface of cells. Its function in the viral replicative cycle is still not fully understood, but it may be involved with the hemagglutinin in promoting fusion of the viral and host cell membranes in the early stages of the infectious cycle (Huang et al., 1980) or prevent aggregation of virus particles during their release from host cells (Palese et al., 1974).

Influenza neuraminidase is a tetramer of four identical glycosylated polypeptides which is coded by one of the eight negative-stranded RNA molecules of the virus. Electron micrographs show that the NA is composed of a square boxlike "head" containing four coplanar subunits and a long thin "stalk" with a small knob at the end (Laver & Valentine, 1969;

† From the Department of Microbiology, John Curtin School of Medical Research, Australian National University, Canberra, A.C.T. 2600 Australia. *Received January 18, 1982.* This work was supported in part by Grant AI 15343 from the National Institute of Allergy and Infectious Diseases.

‡ Present address: Virus Ecology Research Group, Research School of Biological Sciences, Australian National University, Canberra, A.C.T. 2600 Australia.

§ Present address: Department of Microbiology, University of Alabama in Birmingham, Birmingham, AL 35294.

[1] Abbreviations: NA, neuraminidase; NANA, *N*-acetylneuraminic acid; EDTA, ethylenediaminetetraacetic acid; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; RNA, ribonucleic acid; cDNA, complementary deoxyribonucleic acid.

0006-2960/82/0421-4001$01.25/0

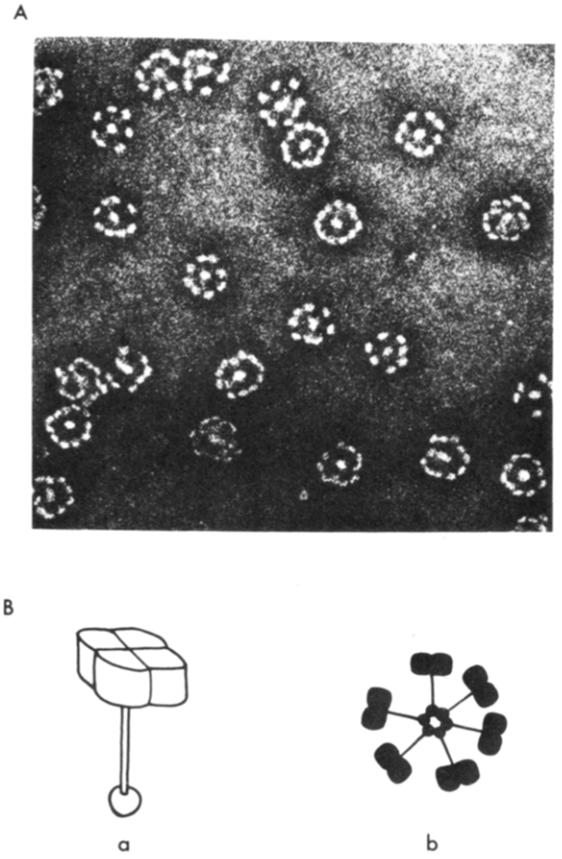

FIGURE 1: (A) Electron micrograph of $NaDodSO_4$-released intact neuraminidase molecules from the recombinant virus A/NWS/33_H-A/Tern/Australia/75_N (H1N9). The $NaDodSO_4$ has been removed, and the NA molecules have aggregated by the hydrophobic tips of their tails which serve to anchor the NA in the lipid of the viral envelope. The average diameter of a rosette is 32 nm. Electron micrograph by N. G. Wrigley. (B) Schematic drawings of (a) the neuraminidase molecule with its "square boxlike head" containing four coplanar subunits and the "stalk" with a hydrophobic knob at the end and (b) a rosette which contains six tetrameric neuraminidase molecules which have "heads" of about 4 × 8 nm and stalks of about 10 nm.

Wrigley et al., 1973; see schematic drawing in Figure 1B). The knob or end of the stalk is embedded in the viral membrane, and isolated NA molecules aggregate via this hydrophobic region (at the base of the stalk) to form rosettes as shown in Figure 1A.

The NA heads which project from the surface can be removed from some strains of influenza virus by digestion with proteases and recovered with complete antigenic and enzymic activity (Mayron et al., 1961; Noll et al., 1962; Wilson & Rafelson, 1963; Drzeniek et al., 1966; Rafelson et al., 1966; Seto et al., 1966; Kendal & Kiley, 1973; Laver, 1978). These heads do not aggregate which indicates that the hydrophobic region of the NA is lost upon protease cleavage, and they also appear by electron microscopy to have lost the thin stalk (Wrigley et al., 1973). In the case of the N2 subtype, Pronase digestion removes 73–75 amino acids from the N terminus of the NA polypeptide (Blok et al., 1982).

Among type A influenza viruses isolated from people, lower mammals, and birds, nine serologically distinct NA subtypes designated N1–N9 have been described (*Memorandum* WHO, 1980). cDNA copies of the RNA segments coding for the NA from two of these subtypes have been sequenced and protein sequences predicted [Fields et al. (1981) and Hiti & Nayak (1982), N1; Markoff & Lai (1982), N2]. Amino acid sequence studies of the isolated intact neuraminidase show that protein synthesis does start at the first AUG codon and that the N terminus of the NA polypeptide is not modified following its synthesis; neither the initiating methionine nor a signal peptide is cleaved off (Blok et al., 1982).

Although the antigenic activity is confined to the heads, a comparison of the serologically unrelated N1 and N2 predicted amino acid sequences (Fields et al., 1981, Hiti & Nayak, 1982; Markoff & Lai, 1982) shows rather more homology in the head region than in the N-terminal portion which is removed by proteases. This N-terminal tail includes the only candidate for a transmembrane sequence and is therefore of structural importance. Since the morphology of many different NA subtypes is similar in the electron microscope (Drzeniek et al., 1968; Laver & Valentine, 1969; Wrigley et al., 1973) we were interested in extending the sequence comparison of the N-terminal region of the NA to the other subtypes and in seeing if the varying sequences could be accommodated into a common structure.

This paper describes the cDNA sequences (15–20% of the complete gene) obtained by priming from the 3′ end of the NA genes from eight of the nine NA subtypes. These sequences contain the coding region for the N terminus of the protein, the transmembrane segment, and the thin stalk and are not involved in antigenic variation. These studies reveal that there is a common sequence at the N terminus of influenza type A neuraminidases, but following this, no amino acids (such as Cys residues) align until around amino acid 85 from the N terminus. There are, however, regions of similar amino acid character which lead to a model for the structure of influenza neuraminidase that is common to all subtypes.

Materials and Methods

Viruses. The influenza virus strains used were A/Memphis/102/72 (H3N2), A/turkey/Oregon/71 (H7N3), A/turkey/Ontario/6118/68 (H8N4), A/Shearwater/Australia/72 (H6N5), A/duck/Alberta/28/76 (H4N6), A/chick/Germany/49 (H10N7), A/Black Duck/Australia/702/78 (H3N8), and one laboratory recombinant, A/Memphis/$1/71_H$-A/Bellamy/42_N (H2N1). The viruses were grown in the allantoic sac of 11-day-old embryonated chick eggs and purified by absorption to and elution from chicken erythrocytes followed by sucrose density gradient centrifugation (Laver, 1969).

Sequencing Procedures. Viral RNA was extracted from purified virus, and the RNA gene segments were separated on 3% polyacrylamide gels containing urea and eluted (Both & Air, 1979). Segment 6 (NA gene) was sequenced by the dideoxy method (Sanger et al., 1977) using the dodecanucleotide d($AGCA_4GCAG_2$) as a primer (Blok & Air, 1980).

Results

The neuraminidase gene from at least one strain from eight of the nine serologically distinct NA subtypes (N1–N8) has been sequenced from the 3′ end, but attempts to sequence the NA genes of several strains from the N9 subtype using the dodecanucleotide primer and the dideoxy method have failed, the most likely explanation being a sequence alteration such that the dodecamer does not prime cDNA synthesis.

The nucleotide and predicted amino acid sequences of the cDNA transcribed from the NA genes of one representative of each of the eight NA subtypes are shown in Figure 2. Peptide data from N2, N5, and N8 strains have shown that the first ATG corresponds to the initiating AUG and that no processing occurs at the N terminus (Blok et al., 1982). These sequences range from 254 to 341 nucleotides in length, which represent 17–23% of RNA segment 6 (NA gene). The first

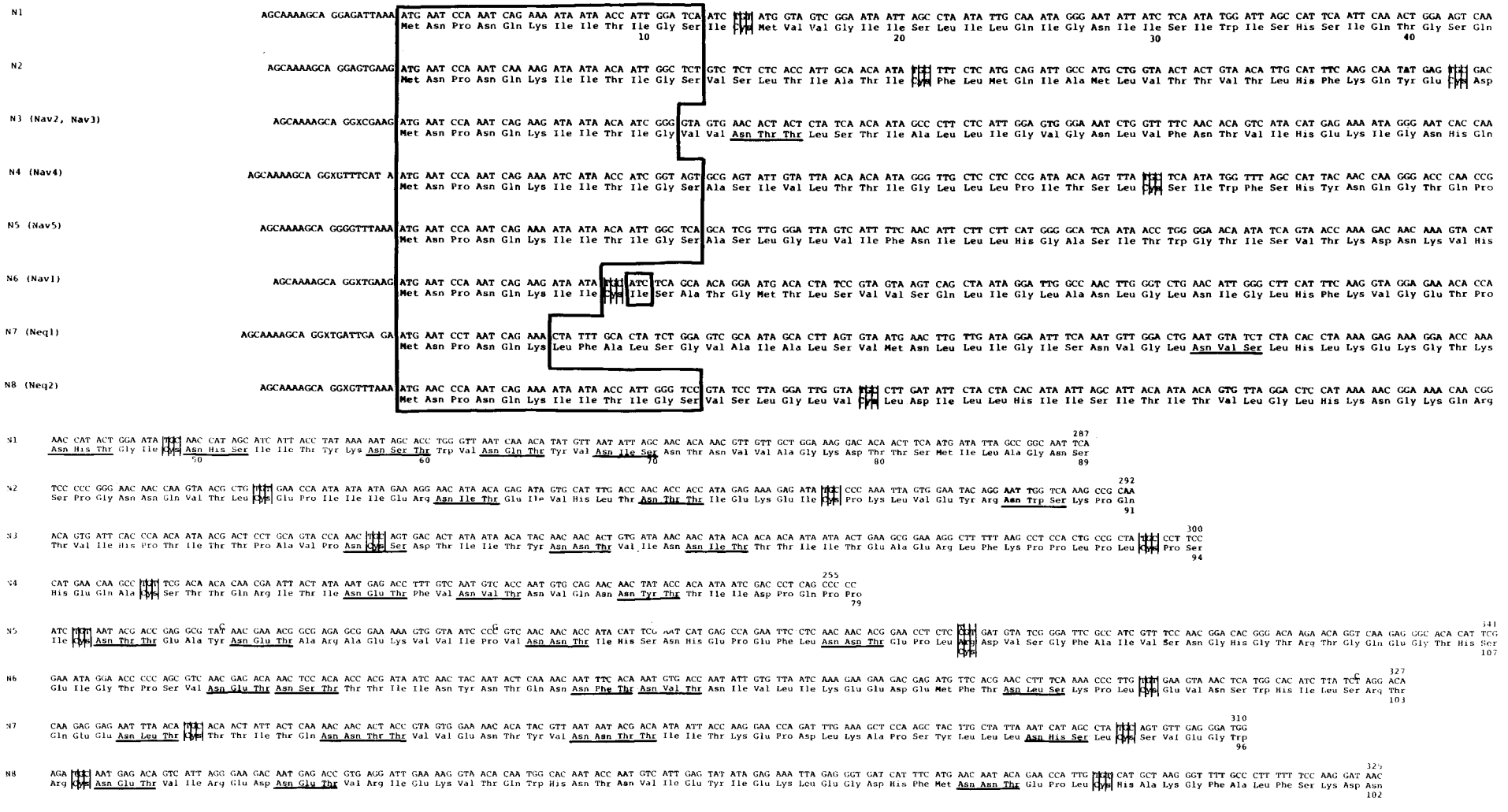

FIGURE 2: Nucleotide and predicted amino acid sequences of influenza neuraminidase from one strain of each of eight NA subtypes, N1–N8: A/Memphis/1/71$_H$–A/Bellamy/42$_N$ (H3N1), A/Memphis/102/72 (H3N2), A/turkey/Oregon/71 (H7N3), A/turkey/Ontario/6118/68 (H8N4), A/Shearwater/Australia/72 (H6N5), A/duck/Alberta/28/76 (H4N6), A/chick/Germany/49 (H10N7), and A/Black Duck/Australia/702/78 (H3N8). The conserved N-terminal region is boxed, while the Cys residues are marked with vertical lines, and the potential glycosylation sites, Asn-X-Thr or Asn-X-Ser, are underlined. The N2 sequence is revised from Blok & Air (1980) by a deletion of one nucleotide at position 161. The possibility of an error in earlier sequences was raised by the complete sequence of A/Udorn/72 (Markoff & Lai, 1982), and although reexamination of most earlier sequencing gels gave the same results, new experiments clearly showed the absence of the extra nucleotide in A/Memphis/102/72.

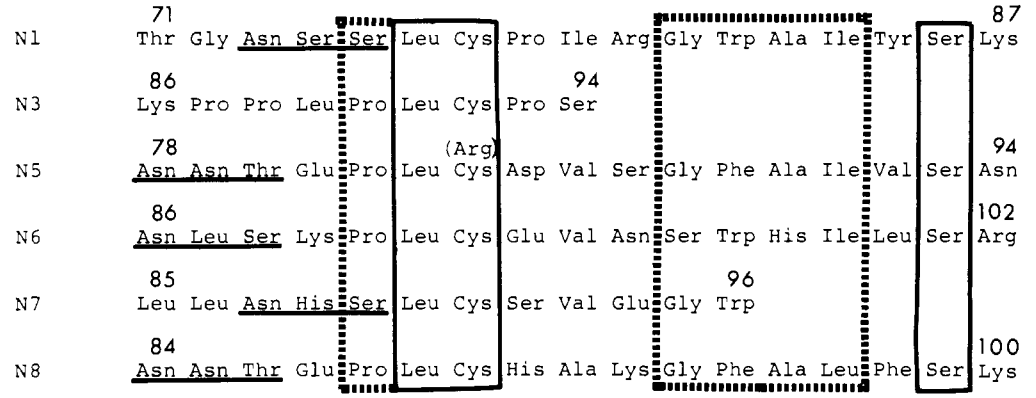

FIGURE 3: Alignment of a Cys residue in the N1 [A/PR/8/34 from Fields et al. (1981)], N3 (A/turkey/Oregon/71), N5 (A/Shearwater/Australia/72), N6 (A/duck/Alberta/28/76), N7 (A/chick/Germany/49), and N8 (A/Black Duck/Australia/702/78) subtypes. The amino acids boxed with solid lines are the same in all of these subtypes while those boxed with dotted lines show two alternative but similar amino acids. The potential glycosylation sites are underlined.

striking feature is that the N-terminal six amino acids of all of the NA subtypes sequences are identical, and this is reflected in the nucleotide sequences. The next six amino acids are the same in five of the eight subtypes, the exceptions being the N3 subtype which differs in one amino acid at position 12, the N6 subtype which has three amino acid differences at positions 9, 11, and 12, and the N7 subtype which is different in all six amino acids (positions 7–12). After this initial similar region at the N terminus, the predicted amino acid sequences are totally different. The changes among the subtypes cannot be explained by a shift in reading frame because the nucleotide sequences are also very different. This is important to note because when the dideoxy method with the RNA template is used, only one strand is sequenced; thus there is always a possibility of a "frameshift" error. Comparison of the amino acid sequences translated from different reading frames in the eight NA subtypes by computer diagonal matching (Gibbs & McIntyre, 1970) reveals no greater homology than that already observed in the correct reading frames.

There are many potential glycosylation sites (Asn-X-Thr/Ser; Neuberger et al., 1972) in the sequences, but they cannot readily be aligned in the various subtypes, although most of them occur in the region 40–80 amino acid from the N terminus. The number of glycosylation sites varies from two (in the N2 and N8 subtypes) to five (in most strains from the N1 subtype) while the most common number is three (in the N3, N4, and N7 subtypes).

All NA subtypes contain at least one Cys residue in the first 75 amino acids from the N terminus, and several subtypes contain more than one cysteine. The various Cys residues in sequences from each subtype were aligned to search for homology among the sequences similar to that seen for the HA, but no similarity in the surrounding sequences could be seen among the different NA subtypes in the N-terminal 80 amino acids. There is, however, a Cys residue near the end of the sequence data which has a common environment in all sequences known. Figure 3 shows the sequences of the N1 (Fields et al., 1981), N3, N5, N6, N7, and N8 subtypes with these Cys residues aligned. Although not all of the amino acid residues around this Cys are the same, several positions contain amino acids which are similar in size or character.

Discussion

In the gene sequences coding for the N-terminal region of the 12 hemagglutinin subtypes, there are certain amino acids, in particular Cys residues, that are conserved throughout the subtypes (Air, 1981). These studies have also revealed that although potential glycosylation sites are not strictly conserved throughout the subtypes, they tend to be present in particular areas of the sequence. It seems that the three-dimensional structure of the HA, stabilized by disulfide bonds, is basically the same in all subtypes, although there is considerable sequence variation (Wilson et al., 1981).

The sequence variation in the N-terminal region of the neuraminidase is enormous. In fact, no homology can be seen, and examination of amino acids 13–80 of the different subtypes does not indicate that they are from proteins with the same enzyme activity.

(A) Model Correlating Sequence Data and the Tetrameric Structure of Influenza Neuraminidase. From electron microscopy studies it is known that the influenza neurminidase is a tetrameric enzyme consisting of a head and a stalk. The hydrophobic region at the base of the stalk serves to anchor the protein in the host-derived lipid membrane of the virus particle. The head carries all the enzyme and antigenic activity. It was not known, however, which part of the polypeptide chain was embedded in the membrane.

From complete nucleotide sequence data, Fields et al. (1981) noted that the only extensive hydrophobic region in the predicted amino acid sequence of PR8 neuraminidase is near the N terminus. The amino acid compositions of tryptic peptides isolated from the intact neuraminidase and Pronase-released neuraminidase heads of the same N2 strain demonstrated that the N terminus is indeed in the tail of the neuraminidase and is removed by Pronase digestion from the heads which contain the C-terminal sequence (Blok et al., 1982). Part of this NA tail serves to anchor the neuraminidase in the viral lipid membrane. The orientation of the influenza neuraminidase is therefore opposite to that of the hemagglutinin, which has its C terminus anchoring the protein in the lipid membrane. Since the gene sequences obtained by priming from the 3′ end of the viral RNA code for the N-terminal region of the NA, the sequence data of the eight NA subtypes shown in Figure 2 represent a structurally important and nonantigenic part of the molecule. The degree of sequence variation is rather surprising.

(1) Hydrophobic Region. Following the N-terminal six amino acids of the predicted protein sequences of the eight NA subtypes, there is a region of 29 amino acids (30 and 31 amino acids in N8 and N5, respectively) which is almost entirely hydrophobic. This hydrophobic region is almost certainly the transmembrane section of the influenza neuraminidase. The average width of the lipophilic core of a biological membrane is about 3 nm (Tanford, 1978), which can be spanned by an extended chain of 10 amino acid residues. The hydrophobic region of the influenza neuraminidase

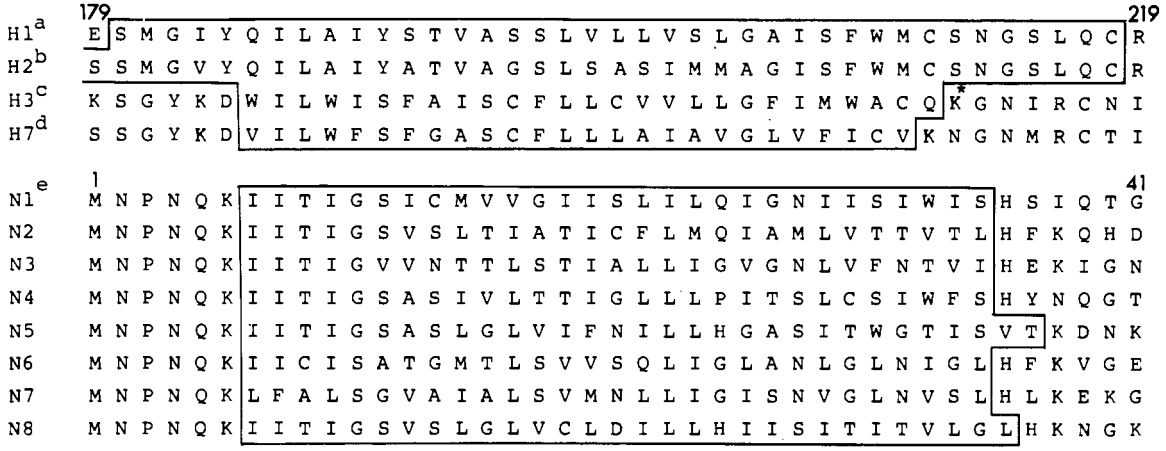

FIGURE 4: Comparison of the hydrophobic regions (boxed) of influenza hemagglutinins (a–d) and neuraminidases (e) from different subtypes. The numbering of amino acids in the HA subtypes start at HA2 residue 1. (a) A/PR/8/34, Winter et al. (1981). (b) A/Japan/305/57, Gething et al. (1980). (c) A/Aichi/2/68 (asterisk = K → R), Verhoeyen et al. (1980); A/Memphis/102/72, Sleigh et al. (1980); A/Victoria/3/75, Min Jou et al., 1980. (d) FPV (fowl plague virus), Porter et al. (1979). (e) For strain designations of N1–N8, see Figure 2.

is 29–31 amino acids long, similar in length to the transmembrane section of the hemagglutinin, and it could form a folded structure. A Chou–Fasman (1974) prediction of secondary structure for the hydrophobic region of the neuraminidase favors the formation of β sheet rather than α helices, but the possibility of forming 3_{10} helices or other coiled structures was not calculated. In one strain (A/Black Duck/Australiia/702/78, N8) there is a charged amino acid residue (Asp) in the hydrophobic region (residue 21). The sequence experiment was repeated several times in case this was a sequencing error, but the same result was always obtained. Folding of the peptide may bury this residue so that the nature of the region is still hydrophobic. It could also loop out of, and then reenter, the membrane.

The hydrophobic sequences of the eight NA subtype proteins are identical for the first six amino acids (Ile-Ile-Thr-Ile-Gly-Ser) in five of the eight NA subtypes, but the next 23, 24, or 25 amino acids are very different. This variation is greater than that observed in the hydrophobic region of the influenza hemagglutinins (Porter et al., 1979; Gething et al., 1980; Min Jou et al., 1980; Sleigh et al., 1980; Verhoeyen et al., 1980; Winter et al., 1981) (see Figure 4), as well as that found in the C-terminal hydrophobic regions of several different variable surface glycoproteins of *Trypanosoma brucei* (Rice-Ficht et al., 1981). The variation of the neuraminidase hydrophobic region is more reminiscent of the variable sequences of the signal peptide of several hemagglutinin subtypes (Air, 1981), but we have no idea why some membrane-insertion sequences should be more restrained in sequence than others.

The sequence at the N terminus preceding the hydrophobic region is identical in the eight subtypes and relatively polar for six amino acids (Met-Asn-Pro-Asn-Gln-Lys). This short hydrophilic peptide could remain on the inside of the viral membrane, and it may be involved in a specific interaction (with the matrix protein, perhaps?).

On the C-terminal side the hydrophobic region of the neuraminidases ends at a potentially charged amino acid residue, His (which may be charged in a protein, depending on the environment) and/or Lys. The basic residues could interact with the phosphate groups at the membrane surface. The sequences following this His and/or Lys must therefore form the stalk region of the neuraminidase.

(*2*) *Stalk Region.* The stalk of influenza neuraminidase was seen as a very thin fiber (≤1.5-nm diameter) in the electron micrographs of $NaDodSO_4$-released neuraminidase (Figure 1; Laver & Valentine, 1969). Neuraminidase is a tetramer of identical subunits, and four polypeptide chains are therefore present in the stalk. Preliminary X-ray data from crystalline Pronase-released NA heads from A/NWS/33_H–A/Tokyo/67_N (H1N2) laboratory recombinant virus reveal the remnant of the stalk as an electron dense column which is less than 1 nm in diameter (Colman & Laver, 1981). This area cannot accommodate four α helices, and it is most likely that the four polypeptide chains are in an extended or very loosely coiled form. The length of the stalk is estimated to be 10 nm in the $NaDodSO_4$-released neuraminidase from the influenza B strain, B/Lee, which would require at least 30 amino acid residues in the extended form. A Chou–Fasman (1974) prediction of secondary structure in the region of the stalk indicates that α helices cannot be formed in any of the NA subtypes nor is the formation of β sheet favorable.

The tail of the neuraminidase from B/Lee has been shown to contain most of the carbohydrate of the protein (Lazdins et al., 1972). Although similar analyses have not been done for influenza A neuraminidases, the predicted amino acid sequences in the region following the hydrophobic section have several potential glycosylation sites. There are up to five glycosylation sites (Asn-X-Ser or Asn-X-Thr; Neuberger et al., 1972) in a stretch of 50 amino acids. The presence of glucosamine in tryptic peptides from A/Bellamy/42 neuraminidase (Blok & Air, 1982) confirms that at least some of these sites are glycosylated. In this region, 40–90 amino acids from the N terminus, there are also several Cys residues. When the cysteines of the different NA subtypes are aligned, no similarity in the flanking sequences is found, in contrast with the HA sequences (Air, 1981). Since this region is part of the apparently extended chain of the stalk, the precise location of Cys residues may not be important for them to be structurally significant. The tetrameric structure of the neuraminidase allows the cysteines of each identical polypeptide chain to align so that disulfide bonds may form between the two pairs of Cys residues.

Electrophoretic studies of $NaDodSO_4$-released NA or protease-released NA heads, under reducing or nonreducing conditions, indicate that the disulfide bond linkage of monomers to form dimers is in the stalk region of the influenza N2 neuraminidase (Blok et al., 1982). Of the eight subtypes sequenced in the stalk, only N6 does not have a Cys residue while several subtypes have more than one. The presence of

at least two potential glycosylation sites in the otherwise extremely variable stalk region of all of the neuraminidase subtypes sequenced suggests that carbohydrate as well as the disulfide bonds may stabilize the extended polypeptide chains.

It is not known exactly where the globular heads of the neuraminidase begin in our predicted amino acid sequences, but if the Cys residue near the end of our sequence data is aligned in the different NA subtypes, the surrounding sequences show significant similarity (Figure 3). This Cys residue could be in the head region and analogous to the structurally important disulfide bonds in HA1 which are conserved among the different HA subtypes (Air, 1981; Wilson et al., 1981). The position of the Cys residue varies from 45 to 55 amino acids from the charged Lys residue near the membrane surface. If this Cys residue is near the start of the head, the variable position indicates that the length of the stalk may vary and that the average length is probably 50 amino acids. This would measure 18.2 nm if the polypeptide were in the extended form (Pauling et al., 1951) which is longer than the 10 nm estimated for B/Lee neuraminidase in the electron microscope. We already know that the length of the stalk can vary even within a single subtype (N1; Blok & Air, 1982).

(*B*) *Membrane Insertion of Influenza Neuraminidase.* The influenza neuraminidase is anchored in the lipid membrane by the N-terminal region. This orientation is opposite to that of the influenza hemagglutinin (Skehel & Waterfield, 1975) and other viral glycoproteins such as the glycoprotein from vesicular stomatitis virus (Rose et al., 1980) and the E_1 and E_2 protein from Semliki Forest virus (Garoff & Söderlund, 1978). There are, however, a few proteins which have their N-terminal regions embedded in the lipid bilayer as does the influenza neuraminidase, for example, the hydrolases of the small intestinal brush border membrane: aminopeptidase (Maroux & Louvard, 1976) and isomaltase (Maroux & Louvard, 1976; Frank et al., 1978; Brunner et al., 1979).

There are two basic views of membrane insertion or translocation of glycoproteins. (1) The signal hypothesis (Blobel & Dobberstein, 1975), postulates the existence of specific membrane proteins which recognize the N-terminal hydrophobic leader sequence or signal peptide and aggregate to form a hydrophilic tunnel. The protein is then translocated across the membrane (through this tunnel) during synthesis followed by endoproteolytic removal of the signal peptide. (2) The other view assumes that there is no specific mechanism involving membrane proteins but that the protein is inserted spontaneously (Bretscher, 1973; Wickner, 1979). A specific example of this view is the loop model (Inouye & Halegoua, 1980). As the sequence extends, a loop is formed while the basic residue which follows the initiating methionine of many secretory or transmembrane proteins remains on the inside surface of the membrane. Exposure of a particular site between the hydrophobic peptide and the mature protein is followed by cleavage at this site.

The influenza hemagglutinin is a typical membrane-bound glycoprotein in that it is synthesized as a precursor protein containing an N-terminal hydrophobic (signal) sequence (Air, 1979; Elder et al., 1979) and that the C-terminal region is embedded in the membrane (Skehel & Waterfield, 1975). It is not known which, if either, of the two basic views is correct for the insertiion of proteins into or transport across membranes, but sequence data from the hemagglutinin are compatible with the "signal hypothesis" since a leader peptide is cleaved off and the mature protein is anchored by its C-terminal region. The influenza neuraminidase, on the other hand, has its N-terminal region embedded in the membrane and is not synthesized as a longer precursor protein. The first six amino acids at the N terminus of the neuraminidase are polar, and one is a basic residue. The neuraminidase could be inserted into the membrane by the loop model, with its six polar residues remaining on the inside surface of the membrane and its N-terminal hydrophobic sequence spanning the lipid bilayer. It has been suggested by Frank et al. (1978) that the isomaltase has an "extended signal sequence" which transfers the protein across the membrane and then remains attached to the membrane. Engelman & Steitz (1981) have proposed a similar mechanism to the loop model, but there is no requirement for basic groups near the N terminus and the model can accommodate membrane-bound proteins with or without a signal peptide, attached by N- or C-terminal hydrophobic sequences. This hypothesis would therefore accommodate both of the influenza surface glycoproteins, hemagglutinin and neuraminidase.

Acknowledgments

We thank Drs. P. Colman, A. Gibbs, G. Laver, C. Ward, D. Wiley, and I. Wilson for helpful discussions, Dr. L. Markoff for providing his N2 sequence prior to publication, and Anne Mackenzie and Kathy Fry for expert technical assistance.

References

Air, G. M. (1979) *Virology 97*, 468–472.
Air, G. M. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 7639–7643.
Blobel, G., & Dobberstein, B. (1975) *J. Cell Biol. 67*, 835–851.
Blok, J., & Air, G. M. (1980) *Virology 107*, 50–60.
Blok, J., & Air, G. M. (1982) *Virology 118*, 229–234.
Blok, J., Air, G. M., Laver, W. G., Ward, C. W., Lilley, G. G., Woods, E. F., Roxburgh, C. M., & Inglis, A. S. (1982) *Virology 119*, 109–121.
Both, G. W., & Air, G. M. (1979) *Eur. J. Biochem. 96*, 363–372.
Bretscher, M. S. (1973) *Science (Washington, D.C.) 181*, 622–629.
Brunner, J., Hauser, H., Braun, H., Wilson, K. J., Wacker, H., O'Neill, B., & Semenza, G. (1979) *J. Biol. Chem. 254*, 1821–1828.
Chou, P. Y., & Fasman, G. D. (1974) *Biochemistry 13*, 222–245.
Colman, P. M., & Laver, W. G. (1981) *Struct. Aspects Recognit. Assem. Biol. Macromol., Proc. Aharon Katzir-Katchalsky Conf., 7th*, 869–872.
Drzeniek, R., Seto, J. T., & Rott, R. (1966) *Biochim. Biophys. Acta 128*, 547–558.
Drzeniek, R., Frank, H., & Rott, R. (1968) *Virology 36*, 703–707.
Elder, K. T., Bye, J. M., Skehel, J. J., Waterfield, M. D., & Smith, A. E. (1979) *Virology 95*, 343–350.
Engelman, D. M., & Steitz, T. A. (1981) *Cell (Cambridge, Mass.) 23*, 411–422.
Fields, S., Winter, G., & Brownlee, G. G. (1981) *Nature (London) 290*, 213–217.
Frank, G., Brunner, J., Hauser, H., Wacker, H., Semenza, G., & Zuber, H. (1978) *FEBS Lett. 96*, 183–188.
Garoff, H., & Söderlund, H. (1978) *J. Mol. Biol. 124*, 535–549.
Gething, M. J., Bye, J., Skehel, J., & Waterfield, M. (1980) *Nature (London) 287*, 301–306.
Gibbs, A. J., & McIntyre, G. A. (1970) *Eur. J. Biochem. 16*, 1–11.
Hiti, A. L., & Nayak, D. P. (1982) *J. Virol. 41*, 730–734.

Huang, R. T. C., Rott, R., Wahn, K., Klenk, H. D., & Kohama, T. (1980) *Virology 107*, 313–319.
Inouye, M., & Halegoua, S. (1980) *CRC Crit. Rev. Biochem. 7*, 339–371.
Kendal, A. P., & Kiley, M. P. (1973) *J. Virol. 12*, 1482–1490.
Laver, W. G. (1969) in *Fundamental Techniques in Virology* (Habel, K., & Salzman, N. P., Eds.) pp 82–86, Academic Press, New York.
Laver, W. G. (1978) *Virology 86*, 78–87.
Laver, W. G., & Valentine, R. C. (1969) *Virology 38*, 105–119.
Lazdins, I., Haslam, E. A., & White, D. O. (1972) *Virology 49*, 758–765.
Markoff, L., & Lai, C. J. (1982) *Virology 119*, 288–297.
Maroux, S., & Louvard, D. (1976) *Biochim. Biophys. Acta 419*, 189–195.
Mayron, L. W., Robert, B., Winzler, R. J., & Rafelson, M. E. (1961) *Arch. Biochem. Biophys. 92*, 475–483.
Memorandum WHO (1980) *Bull. W.H.O. 58*, 585–591.
Min Jou, W., Verhoeyen, M., Devos, R., Saman, E., Fang, R., Huylebroeck, D., & Fiers, W. (1980) *Cell (Cambridge, Mass.) 19*, 683–696.
Neuberger, A., Gottschalk, A., Marshall, R. D., & Spiro, R. G. (1972) in *The Glycoproteins: Their Composition, Structure and Function* (Gottschalk, A., Ed.) pp 450–490, Elsevier, Amsterdam.
Noll, H., Aoyagi, R., & Orlando, J. (1962) *Virology 18*, 154–157.
Palese, P., Tobita, K., Ueda, M., & Compans, R. W. (1974) *Virology 61*, 397–410.
Pauling, L., Corey, R. B., & Branson, H. R. (1951) *Proc. Natl. Acad. Sci. U.S.A. 37*, 205–211.
Porter, A. G., Barber, C., Carey, N. H., Hallewell, R. A., Threlfall, G., & Emtage, J. S. (1979) *Nature (London) 282*, 471–477.
Rafelson, M. E., Gold, S., & Priede, I. (1966) *Methods Enzymol. 8*, 677–680.
Rice-Ficht, A. C., Chen, K. K., & Donelson, J. E. (1981) *Nature (London) 294*, 53–57.
Rose, J. K., Welch, W. J., Sefton, B. M., Esch, F. S., & Ling, N. C. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 3884–3888.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Seto, J. T., Drzeniek, R., & Rott, R. (1966) *Biochim. Biophys. Acta 113*, 402–404.
Skehel, J. J., & Waterfield, M. D. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 93–97.
Sleigh, M. J., Both, G. W., Brownlee, G. G. Bender, V. J., & Moss, B. A. (1980) in *Structure and Variation in Influenza Virus* (Laver, W. G., & Air, G. M., Eds.) pp 69–80, Elsevier, New York.
Tanford, C. (1978) *Science (Washington, D.C.) 200*, 1012–1018.
Verhoeyen, M., Fang, R., Min Jou, W., Devos, R., Huylebroeck, D., Saman, E., & Fiers, W. (1980) *Nature (London) 286*, 771–776.
Wickner, W. (1979) *Annu. Rev. Biochem. 48*, 23–45.
Wilson, I. A., Skehel, J. J., & Wiley, D. C. (1981) *Nature (London) 289*, 366–373.
Wilson, V. W., & Rafelson, M. E. (1963) *Biochem. Prep. 10*, 113–117.
Winter, G., Fields, S., & Brownlee, G. G. (1981) *Nature (London) 292*, 72–75.
Wrigley, N. G., Skehel, J. J., Charlwood, P. A., & Brand, C. M. (1973) *Virology 51*, 525–529.

Role of Poly(adenosine diphosphate ribose) in Deoxyribonucleic Acid Repair in Human Fibroblasts†

Michael R. James and Alan R. Lehmann*

ABSTRACT: We have investigated the role of poly(adenosine diphosphate ribose) in DNA repair in human fibroblasts by observing the effects of 3-aminobenzamide (3AB), a specific inhibitor of poly(ADP-ribose) synthesis, on various aspects of DNA repair. After treatment of human fibroblasts with dimethyl sulfate (DMS), 3AB retarded the joining of strand breaks; unscheduled DNA synthesis was unaffected after low doses of DMS but was stimulated after high doses. 3AB also enhanced the cytotoxicity of DMS. After γ irradiation there was a slight inhibition by 3AB of the rejoining of single-strand breaks but no effect on the rejoining of double-strand breaks, unscheduled DNA synthesis, DNA replicative synthesis, or cytotoxicity. There were no effects of 3AB on the repair of UV damage. On the basis of the different kinetics of the various steps of excision repair processes after different treatments of fibroblasts, our results are interpreted as evidence that the synthesis of poly(ADP-ribose) is involved in the ligation step of excision repair.

ADP-ribosylation is the covalent addition to protein of the ADP-ribose portion of an NAD molecule. ADP-ribosyltransferase (ADPRT; EC 2.4.99)[1] is a DNA-dependent nuclear enzyme which catalyzes both the ADP-ribosylation of chromosomal proteins and the further addition of ADP-ribose moieties to produce a homopolymer of up to 70 residues (Hayaishi & Ueda, 1977; Butt & Smulson, 1980; Purnell et

† From the Medical Research Council, Cell Mutation Unit, University of Sussex, Falmer, Brighton, BN1 9QG, United Kingdom. *Received December 29, 1981; revised manuscript received May 4, 1982.* This work was supported in part by EC Contract 166-76-1-BIO-UK.

[1] Abbreviations: 3AB, 3-aminobenzamide; ADPRT, adenosine diphosphoribosyltransferase; DMS, dimethyl sulfate; DSBs, double-strand breaks; MNU, *N*-methyl-*N*-nitrosourea; NAD, nicotinamide adenine dinucleotide; PBS, phosphate-buffered saline; SSBs, single-strand breaks; UDS, unscheduled DNA synthesis; UV, shortwave ultraviolet light; Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; EDTA, ethylenediaminetetraacetic acid.

0006-2960/82/0421-4007$01.25/0