JOURNAL OF VIROLOGY, Oct. 1988, p. 3824–3831
0022-538X/88/103824-08$02.00/0

Vol. 62, No. 10

Redundancy of Signal and Anchor Functions in the NH_2-Terminal Uncharged Region of Influenza Virus Neuraminidase, a Class II Membrane Glycoprotein

DONALD J. BROWN, BRENDA G. HOGUE, AND DEBI P. NAYAK*

Department of Microbiology and Immunology, University of California at Los Angeles School of Medicine, and Jonsson Comprehensive Cancer Center, Los Angeles, California 90024-1747

Received 20 May 1988/Accepted 7 July 1988

Class II membrane glycoproteins share a common topology of the NH_2 terminus inside and the COOH terminus outside the cell. Their transport to the cell surface is initiated by the function of a single hydrophobic domain near the NH_2 terminus. This functional domain serves both as an uncleaved signal sequence and as a transmembrane anchor. We examined the signal and anchor functions of influenza virus neuraminidase, a prototype class II membrane glycoprotein, by deletion analysis of its long, uncharged amino-terminal region. The results presented here show that the entire stretch of 29 uncharged amino acids (7 to 35) is not required for either a signal sequence or an anchor sequence function. On the basis of translocation and membrane stability data for different mutants, we suggest that the first 20 amino acid residues (7 to 27) are likely to provide the hydrophobic core for these functions and that within this putative subdomain some sequences are more efficient than the other sequences in providing a translocation function. Finally, it appears that neuraminidase and its mutant proteins are translocated with the proper orientation, regardless of the characteristics of the flanking sequences.

Recently, a great deal of attention has focused on sequences which direct proteins to their subcellular compartments. Perhaps the best studied is the signal sequence which directs a protein to the rough endoplasmic reticulum (RER) and is involved in translocation of the protein across the lipid bilayer (37). The common characteristics of these sequences are their overall hydrophobicity with a minimum core of eight or nine apolar amino acids (31, 33). It is not surprising that a great degree of heterogeneity is found among signal sequences (38). In fact, a significant portion of randomly obtained sequences provides signal information (17). Additionally, other sequences not typically involved in translocation, such as stop transfer-anchor domains, substitute for a signal sequence when they are engineered to the proper amino-terminal location (23, 40).

Whereas most signal sequences are removed by proteolysis from the mature protein (37), some proteins contain hydrophobic leaders which are not cleaved and are also used as the anchor domain for the mature protein (21, 29, 39, 41). Since both of these functions are determined by hydrophobic amino acids, and anchor sequences are typically longer than signal sequences, it has been postulated that the signal function is redundant within the anchoring domain. This was demonstrated to be true for the asialoglycoprotein receptor. In this case, a minimal length of hydrophobic residues was found to provide the signal function, whereas the additional hydrophobic residues were necessary to ensure proper anchoring (28). These data strongly suggest that, in this case, no subdomain structure is required for a signal-anchor domain.

Influenza virus neuraminidase (NA), a class II membrane glycoprotein, contains a signal-anchor domain within the first 40 amino acids of the mature protein (3). Located within this domain is an uncharged hydrophobic region of 29 amino acids (6). Previous results from this laboratory, using site-directed mutagenesis to place charged residues within this region (27), suggested that there is a specific subdomain within the larger anchor region which is required for the RER-targeting signal. However, the presence of a charged amino acid may not always cause disruption of the signal function of the adjacent sequences (8). Therefore, in the present study we created a nested set of deletion mutants within this region. These mutants were then tested with an in vitro transcription-translation system for retention of a functional signal and for ability to be anchored in the membrane.

MATERIALS AND METHODS

Plasmid construction. Various construction schemes used to generate the deletion mutants are described below. Wild-type (wt) NA cDNA (1,465 base pairs) of A/WSN/33 influenza virus (H1N1) was removed from SNC (9, 15) with *Eco*RI and directly subcloned into pGem3 (Promega Biotech) to create pNA. pN15 was created from SN10 (9), a clone generated by BAL 31 digestion of wt NA cDNA and subcloned into pA11SVL3 after addition of *Eco*RI linkers. To construct pN15, the NA sequences of SN10 were removed with *Eco*RI and subcloned into the *Eco*RI site of pGem3. This mutant (pN15) begins translation at the first available AUG (amino acid 15 of wt NA). pN18 was created by restricting SN11 (27) with *Acc*I (nucleotide 66 of SN11), filling in the overhangs with the Klenow fragment, and then recutting with *Eco*RI. The fragment representing bases 66 to 1,400 of wt NA was then cloned into pGem3, which had been cut with *Acc*I, filled in, and then recut with *Eco*RI. This mutant (pN18) relies on an AUG provided in the polylinker of the vector and adds six amino acids to the authentic NA sequence in front of residue 18. To generate pN23, SNC (9) was restricted with *Ssp*I (nucleotide 88) and *Eco*RI, the DNA fragment representing bases 88 to 1,400 of wt NA, was then cloned into pGem3 treated as described for pN18. This mutant (pN23) also uses the AUG of the vector and adds five

* Corresponding author.

```
                  7                                35     40 aa
Wild Type   MNPNQK IITIGSICMVVGIISLILQIGNIISIWIS HSIQTG
N15                      MVVGIISLILQIGNIISIWIS HSIQTG
N18                  MPAGRD GIISLILQIGNIISIWIS HSIQTG
N23                      MPAGR LQIGNIISIWIS HSIQTG
N26                      MVLRPDPNS GNIISIWIS HSIQTG
N23-122     MNPNQK IITIGSICMVVGIISLI\      /ALLN DKHSRG
N17-23      MNPNQK IITIGSICMV\     /LQIGNIISIWIS HSIQTG
N17-26      MNPNQK IITIGSICMV\     /ANSGNIISIWIS HSIQTG
```

FIG. 1. Predicted amino acid (aa) sequences of the NH_2 termini of wt and mutant NA proteins. The plasmids shown were sequenced by chain termination (26) and read through the appropriate junctions. The bold sequence indicates the hydrophobic region as predicted by the Kyte and Doolittle method (18). The underlined residues are those derived from a non-NA sequence as described in Materials and Methods. The symbols \ and / indicate the internally deleted sequences of NA.

amino acids in front of residue 24 of wt NA. pN26 was derived from SN26 (9). Here, SN26 was restricted with *Kpn*I (a site within simian virus 40 and 60 nucleotides upstream of the NA sequences) and *Hin*dIII (a site downstream of NA in simian virus 40). This DNA fragment was then cloned into the polylinker of pGem3 between these sites. pN26 uses the simian virus 40 agno gene AUG codon and adds nine amino acids in front of residue 27 of wt NA (9). pN23-122 was generated by ligation of two fragments obtained from pNA. One fragment was obtained by restricting pNA (described above) with *Ssp*I (nucleotide 88 of NA). The other *Ssp*I fragment was recut with *Nar*I (nucleotide 385 of NA) and filled in with the Klenow fragment. Religation of these two fragments removed the sequences of NA which encode amino acids 23 to 122 and produced pN23-122. To construct pN17-23, two fragments of pN11 were used. pN11 was created by cloning the NA-containing sequences of SN11 into pGem3 at the *Eco*RI site. pN11 was restricted with *Acc*I (nucleotide 66 of NA), and the overhangs were removed with mung bean nuclease and recut with *Ssp*I (nucleotide 1975 of pGem3). This fragment was isolated and ligated with the large DNA fragment obtained after digestion of pN11 with *Ssp*I (nucleotides 88 of NA and 1975 of pGem3). pN17-23 lacks the coding sequence of amino acids 17 to 23 of wt NA. To create pN17-26, pN11 was restricted with *Acc*I and filled in with the Klenow fragment, and the large fragment (pGem3 plus 66 nucleotides of NA) was isolated. This fragment was ligated to a *Bam*HI fragment obtained from SN26 blunted with mung bean nuclease (bases −6 to 1098 of NA). pN17-26 lacks the coding sequence for amino acids 17 to 26 and is also truncated at the 3′ end by ≈300 nucleotides. All manipulations were performed by standard techniques, and the enzymes were obtained from Bethesda Research Laboratories, Inc., or New England BioLabs, Inc.

To confirm the constructions, each mutant was sequenced through the appropriate junction by the chain termination method of Sanger et al. (26). The predicted amino acid sequence for each mutant is shown in Fig. 1.

Transcriptions. Plasmid DNA was purified by the method of Krieg and Melton as described by Promega Biotec. Three micrograms of either linearized or circular DNA, was transcribed in a 20-μl transcription mixture composed of 40 mM Tris hydrochloride (pH 8.3), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM dithiothreitol, 20 μg of human placental RNase inhibitor (isolated as described by Blackburn et al. [2]) per ml, 0.1 mg of nuclease-free bovine serum albumin (Boehringer Mannheim Biochemicals), 1 mM each ATP, CTP, and UTP, 0.05 mM GTP, 1 mM GpppG (Pharmacia), and 10 U of either T7 or Sp6 RNA polymerase. The reaction was performed at 39°C for 20 min and then supplemented with GTP to a final concentration of 1 mM and allowed to react for another 20 min. At the end of the incubation period, the transcription mixture was used immediately for translation or quick-frozen in liquid nitrogen for storage at −70°C.

Translations. Five microliters of the transcription mixture was added to a translation cocktail composed of 35 μl of rabbit reticulocyte lysate (Promega Biotec), 1 μl of amino acids without methionine (1 mM), 5 μl of [^{35}S]methionine (15 mCi/ml; Amersham Corp.), 2.5 μl of RNase inhibitor, 1 μl of aprotinin (10 mg/ml; Boehringer Mannheim Biochemicals), and 5 μl of nuclease-treated (35) dog pancreas microsomes (RM) or RM buffer (50 mM triethanolamine [pH 7.5], 250 mM sucrose, 1 mM dithiothreitol). This was incubated at 30°C for 45 min and then supplemented with 1 μl of 1 mM methionine and incubated for a further 15 min.

At the end of the incubation period, a 10-μl sample of the cocktail was removed, adjusted to 1% sodium dodecyl sulfate (SDS), and boiled for 3 min. This was diluted 10-fold into 50 mM phosphate buffer (pH 5.5) and split into two aliquots. One aliquot was supplemented with 8 mU of endoglycosidase H (endo H; ICN Pharmaceuticals Inc.; 30), and both were incubated overnight at 37°C. On the next day, the samples were adjusted to 2.3% SDS–50 mM Tris hydrochloride (pH 6.8)–10% glycerol. The samples were boiled for 3 min and analyzed by SDS-polyacrylamide gel electrophoresis on 10% gels (19).

Another 10-μl sample of the translation cocktail was removed and adjusted with RSB (10 mM Tris hydrochloride [pH 7.4], 10 mM KCl, 1.5 mM $MgCl_2$), 2 μl of 100 mM $CaCl_2$, and 2 μl of 30 mM tetracaine (24) to a final volume of 20 μl. This was split into three 6-μl samples, one receiving an additional 1 μl of trypsin (1.25 mg/ml) and another receiving trypsin and 1 μl of 10% Triton X-100. The samples were incubated on ice for 30 min and then immediately boiled in SDS.

As a measure of translocation efficiency, 10 μl of the cocktail was diluted into 100 μl of ice-cold RSB, incubated for 30 min on ice, and then spun in a microcentrifuge for 30 min to pellet the RMs or any particulate material. The supernatant fraction was precipitated with 10 volumes of ice-cold acetone, recovered by centrifugation, and prepared for electrophoresis. The pellets were carefully rinsed with cold water and boiled in SDS.

Carbonate extractions (14) were performed by diluting 10 μl of the translation cocktail into 100 μl of 100 mM sodium carbonate (pH 10 or 11) and incubating the mixture on ice for 30 min. The samples were then centrifuged for 30 min in an Eppendorf microcentrifuge. The soluble proteins were recovered by acetone precipitation and boiled in SDS. The pellets were rinsed and prepared for electrophoresis. Both the soluble and pellet fractions were analyzed by SDS-polyacrylamide gel electrophoresis.

RESULTS

Figure 1 shows the predicted amino acid sequences of the wt NA and each of the mutants constructed. The wt NA possesses a small cytoplasmic tail composed of 6 amino acids followed by 29 uncharged residues. Previously, we have shown that the first 40 amino acids contains the information necessary for both translocation and anchoring

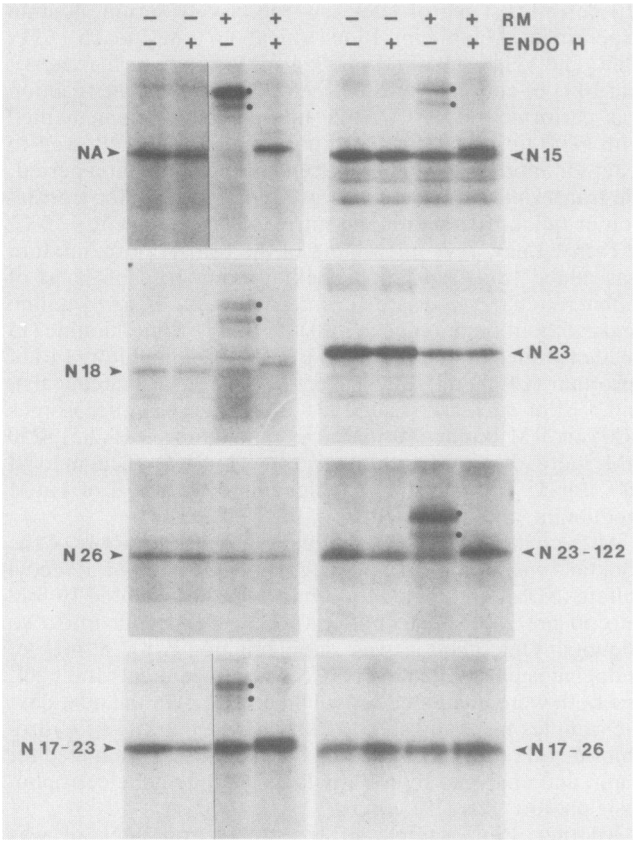

FIG. 2. Translation and glycosylation of NA and mutants. In vitro translations of NA or mutant mRNA in rabbit reticulocyte lysates were performed in the presence (+) or absence (−) of RM. wt and mutant protein products, translated without RM, migrated at 46 kDa (arrowheads), except for N17-26 and N23-122, which gave rise to products of 36 kDa (see Materials and Methods). Dots indicate the higher-molecular-weight protein products made in the presence of RM, presumably representing glycosylation of the four predicted sites, except for N23-122, which has only two glycosylation sites. After endo H treatment (+), the higher-molecular-weight bands disappeared and migrated with or close to the translation products made in the absence of RM (−). The products made in the absence of RM were unaffected by endo H treatment.

of NA (3). Each of the mutants is progressively shortened within the uncharged region of NA (Fig. 1). For convenience, the mutants were designated to indicate the first authentic NA residue retained in the mutant (for amino-terminal deletions) or to indicate the residues removed from NA (for internal deletions). For example, N15 begins at residue 15 of the wt, whereas N17-23 has amino acids 17 through 23 removed. In some cases, extraneous sequences were added because of addition of linkers. These are underlined in Fig. 1.

When the wt NA cDNA was transcribed and translated in rabbit reticulocyte lysates (Fig. 2, NA, −RM, and −endo H), a prominent radioactive band (arrowhead) that migrated at 46 kilodaltons (kDa) was seen. This is consistent with the predicted molecular weight of influenza virus strain WSN NA (15). Additionally, antiserum against strain WSN immunoprecipitated this band (data not shown), confirming that this species is authentic NA. This protein was unaffected in its mobility when treated with endo H, an enzyme which specifically removes N-linked carbohydrates of the high-mannose type (30), a characteristic of localization in the RER. When the wt protein was translated in the presence of RM, two prominent radioactive bands were observed which migrated with an increase in apparent molecular mass (≈55 kDa). These bands (indicated with dots) were of the expected molecular weight for WSN NA glycosylated at three and four of the four predicted glycosylation sites (15). When this sample was treated with endo H, these higher-molecular-weight bands disappeared and a single radioactive band appeared at a position only slightly higher than that of nonglycosylated NA. These results demonstrate that NA became glycosylated in the presence of RMs, presumably because of translocation across the RER into the lumen of the membranes. It should also be noted that after treatment with endo H, the protein migrated in a manner which suggests that the signal peptidase present in the RM failed to cleave wt NA, as in eucaryotic cells.

The behavior of each of the mutants was analyzed in a manner identical to that described above for the wt NA. The results are presented in Fig. 2. In each case, the primary translation product obtained in the absence of RMs (arrowhead) was of the expected molecular weight. The additional band seen just above the N18 and N26 mutant proteins represents a monospecific 48-kDa methionine-binding protein present in the lysates. When translated in the presence of RM, N15, N18, N23-122, and N17-23 produced products which migrated with higher apparent molecular weights (dots). These additional bands were sensitive to the activity of endo H, which caused a shift in their mobility to a position slightly heavier than that of the unglycosylated product. This suggests that these mutants, like the wt protein, were not processed by signal peptidase. In other words, the constructs did not uncover a cryptic cleavage site for the signal peptidase, as reported for another class II membrane protein (20). The other mutants, N23, N26, and N17-26, did not appear to become glycosylated, since they failed to be altered in mobility because of the presence of RM.

Whereas glycosylation of a protein is an indication of translocation, more direct evidence of translocation is protection of the protein against externally added protease. wt NA, translated in the absence of RM (Fig. 3, arrowhead), was completely digested by trypsin (compare lanes 1 and 2). However, in the presence of RM, the glycosylated forms of NA (dots) were resistant to this concentration of trypsin. To demonstrate that glycosylation of the protein did not confer this resistance to protease, a portion of this sample was treated with Triton X-100 to solubilize the microsomes before addition of trypsin. In this case, the glycosylated forms of NA were also completely sensitive to the protease.

The mutants which were glycosylated (N15, N18, and N17-23) demonstrated behavior identical to that of the wt protein (Fig. 3). That is, the nonglycosylated protein was digested, whereas the glycosylated protein was protected, unless the integrity of the microsomes was disrupted by detergent. However, the mutant proteins which did not display any apparent glycosylation (N23, N26, and N17-26) were completely digested by the added trypsin, regardless of the presence or absence of RM or detergent.

As mentioned earlier, the region of NA which functions as a signal sequence also serves as the anchor domain for the mature protein. Therefore, it was important to determine whether these two functions could be dissociated. Samples of the translation cocktail were diluted into cold RSB or carbonate at either pH 10 or pH 11. wt NA, when translated in the absence of RM (Fig. 4, upper left panel), was recovered in the soluble fractions, as expected for a soluble protein. This suggests that NA, although fairly hydrophobic

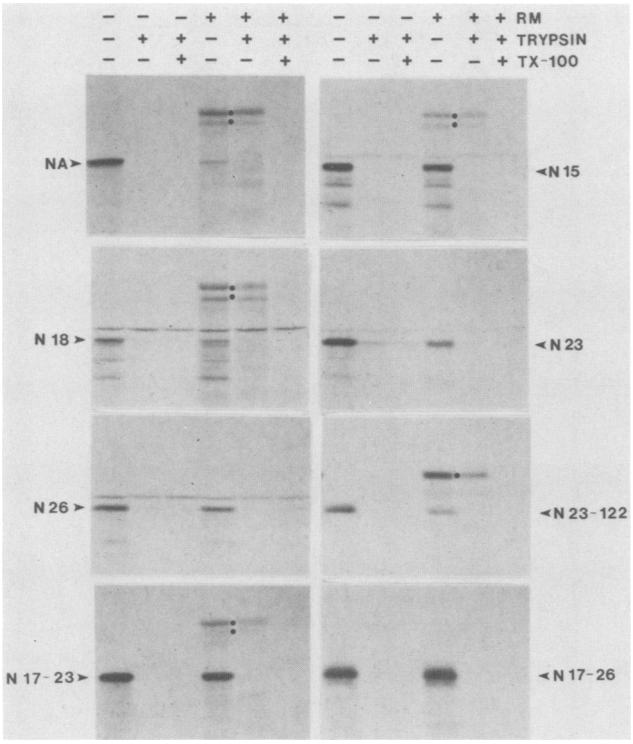

FIG. 3. Protease treatment of translated products. Each transcript was translated in the presence (+) or absence (−) of RM and then divided into three eliquots. The first received neither trypsin (−) nor Triton X-100 (TX-100) (−). The second received 125 μg of trypsin per ml (+) but no Triton X-100 (−). The third received both trypsin (+) and Triton X-100 (+). The glycosylated forms (dots) of NA, N15, N18, N17-23, and N23-122 are shown above the nonglycosylated forms (arrowheads) and were protected from digestion until the RM were disrupted by addition of Triton X-100.

in nature, did not appear to aggregate and become insoluble under these conditions. When these experiments were conducted for NA translated in the presence of RM, virtually all of the protein was recovered in the pellet fraction after dilution into RSB. This suggests that NA was efficiently translocated and that the conditions used did not severely disrupt the integrity of the microsomes. When diluted into carbonate, NA was recovered mostly in the pellet fraction and only a small fraction of the glycosylated protein was found in the supernatant. This indicates that NA was stably anchored in the lipid bilayer.

Although the mutants N23, N26, and N17-26 appeared not to be translocated, on the basis of either glycosylation or protease protection, one could not rule out the possibility that they were translocated in the wrong orientation (i.e., the COOH terminus outside and the NH_2 terminus inside the vesicle; 23, 25). If this were the case, they would neither become glycosylated nor display resistance to protease but should display some association with the microsomal membrane. However, these mutants (N23, N26, and N17-26) were quantitatively recovered in the soluble fractions under all conditions (Fig. 4). This reemphasizes the complete lack of signal function in these mutants.

The primary translation product of N15 (without RM) was always recovered exclusively in the soluble fractions. When translated in the presence of RM and diluted into RSB, the soluble fraction contained virtually all of the nonglycosylated protein and none of the glycosylated protein. The pellet fraction was composed entirely of the glycosylated protein. This demonstrates that translocation of this mutant was directed in a manner similar to that of the wt and none of the protease-sensitive, unglycosylated N15 protein was associated with the RM in a backward orientation. When diluted into carbonate at either pH, the glycosylated forms were recovered in both the soluble and pellet fractions, indicating that this protein, in comparison with the wt, was less stably anchored in the membrane.

N18 gave results similar to that of N15 except that the glycosylated protein was recovered almost exclusively in the pellet fraction at pH 10 while at pH 11 it was found in both fractions. N23-122 behaved in a manner very consistent with the wt NA in that the translocated, glycosylated forms were predominantly recovered with the membrane pellets under these conditions. N17-23 behaved much like N15 in that the translocated protein was recovered as both soluble and anchored species after carbonate extraction at either pH, indicating that it was poorly membrane bound. These results, consistent with the observations of Dalbey and Wickner (8), show that all the mutant proteins translocate in the proper orientation, regardless of the flanking charge (N15 and N18) and each mutant contains a stop transfer function.

Since the wt NA possesses an unusually long uncharged region (29 amino acids) which could satisfy the requirements for a helical hairpin configuration (12), it was of interest to examine the role of the signal recognition particle (SRP) (36) in the translocation of this molecule. We originally attempted to demonstrate translational arrest (34) of NA by exogenous addition of SRP to wheat germ extracts but were unable to observe any effect on translation due to the presence of SRP (data not shown).

These results were unexpected and suggested the possibility that translocation of NA is uniquely independent of the SRP. To assess this possibility, salt-washed, SRP-depleted RM were prepared by washing RM three times in high-potassium buffer (36) to remove the endogenous SRP before use in a standard translocation assay. NA was not glycosylated when translated in the presence of salt-washed, SRP-depleted RM but became glycosylated when the SRP (30 nM) was exogenously added to the salt-washed membranes (Fig. 5). These results demonstrate that NA was dependent upon SRP for translocation across the RER.

DISCUSSION

In previous studies from this laboratory, we demonstrated that the first 40 amino acids of NA contain the functional elements required for RER translocation and transmembrane anchoring (3). A key element required for both of these functions is overall hydrophobicity (13, 31–32, 33). Kyte and Doolittle analysis (10, 18) demonstrated that the length of the hydrophobic domain extends from amino acids 6 to 35 (Fig. 1). This stretch of 29 uncharged residues is longer than the 17 to 20 residues required to span the lipid bilayer of a membrane and remarkably longer than the typical RER translocation signal (10 to 15 residues). By introducing charged amino acids in this region, we suggested that the entire span of 29 amino acids is not required for either the signal or the anchor function (27). Therefore, we undertook this study to further characterize the structural requirements for both signal and anchor functions. Our results are presented in a semiquantitative manner in Table 1.

NA remains anchored in the lipid bilayer of the membrane by the amino-terminal hydrophobic domain with a small

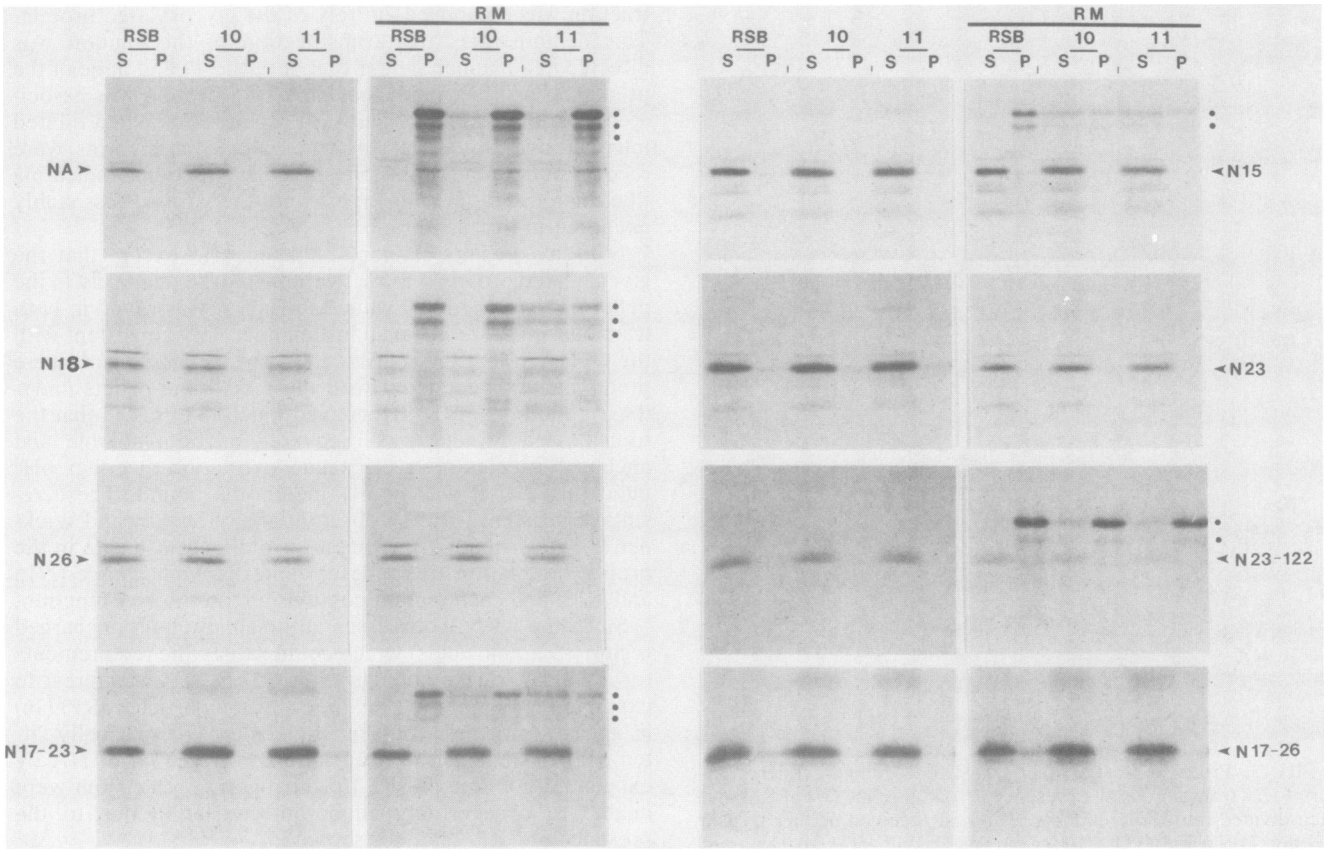

FIG. 4. Membrane stability of translocated proteins. wt NA and translation-positive mutants (N15, N18, N17-23, and N23-122) were translated in the presence (+) or absence of RM. The samples were separated by centrifugation in RSB to indicate the soluble (S) or vesicle-associated (P) fraction. The RM were also subjected to carbonate extraction (14) at pH 10 and pH 11, and the membrane (P) and soluble (S) proteins were isolated as described in Materials and Methods. The arrowheads indicate the nonglycosylated proteins, and the band above these are glycosylated forms (dots).

NH_2-terminal tail exposed in the cytoplasm. The precise boundary of the transmembrane domain is not known, though it has been suggested that it encompasses the 29 uncharged residues from 7 to 35 (6). Although the precise requirements for the anchoring function are not known, it is well accepted that transmembrane anchors are hydrophobic, exist in a helical conformation in the lipid environment, and require approximately 20 amino acid residues in length to span the width of the lipid bilayer (see reference 13 for a review). When the span of hydrophobicity is partially deleted or interrupted by charged amino acids, membrane proteins often become less stably integrated or are sometimes secreted (1, 7).

The data presented here show that we could remove large portions from either end of this region (amino acids 7 to 35) and still retain a reasonably stable membrane interaction, indicating the redundancy of the sequence information for the anchor function within this region. However, our data also suggest that removal of the sequences from the NH_2-terminal (N18 and N15) or central region (N17-23) affects the anchoring function more dramatically than does removal of sequences from the COOH terminus of this region (N23-122). Furthermore, the presence or absence of charges in the amino terminus also appeared to affect membrane stability (N15 versus N18). Overall, our data support the idea suggested previously that residues 27 to 35, which are less hydrophobic and resemble the C region of a typical signal

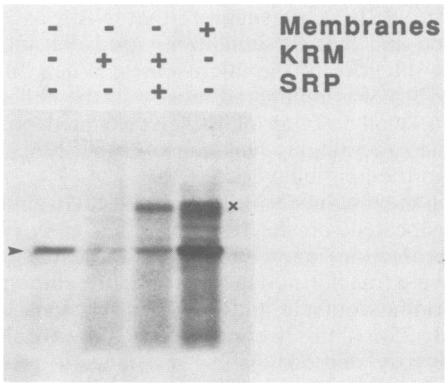

FIG. 5. SRP-dependent translocation of NA. wt NA was translated in wheat germ embryo extracts with (+) or without (−) the SRP and supplemented with (+) or without (−) salt-washed, SRP-depleted microsomes (KRM). NA did not become glycosylated unless translated in the presence of both KRM and SPR. Fully glycosylated NA comigrated with the glycosylated product obtained when NA was translated with nonextracted microsomal membranes (+). Arrowhead, nonglycosylated NA; ×, glycosylated NA.

TABLE 1. Quantitation of the functional behavior of NA and mutant proteins in vitro[a]

Protein	% Glycosylated[b]	% Translocated[c]	% Anchored at[d]:	
			pH 10	pH 11
NA	86	98	75	71
N15	33	49	35	19
N18	68	90	86	33
N23	ND	ND	NA	NA
N26	ND	ND	NA	NA
N23-122	61	83	79	84
N17-23	35	37	37	33
N17-26	ND	ND	NA	NA

[a] Each value listed varied by 10%, depending on the trace used for integration. ND, Not detected; NA, not applicable.

[b] Determined by densitometric tracing of the lanes listed as + RM, − trypsin, − TX-100 in Fig. 3.

[c] Determined by densitometric tracing (Bio-Rad model 620) of the lanes listed as S, RSB versus P, RSB in Fig. 4 after autoradiographic exposure of preflashed Kodak XAR-5 film.

[d] Determined by densitometric tracing of the fully glycosylated protein in S versus P of Fig. 4 for the carbonate-extracted samples.

	7	27	36 aa	SIGNAL
NA	MNPNQK	IITIGSICMVVGIISLILQIG	NIISIWISHSIQTG	+
N10*	MVLRPDPNS	GSICMVVGIISLILQIG	NIISIWISHSIQTG	+
N15		MVVGIISLILQIG	NIISIWISHSIQTG	+
N18		MPAGRD GIISLILQIG	NIISIWISHSIQTG	+
N23		MPAGR LQIG	NIISIWISHSIQTG	-
N26		MVLRPDPNS G	NIISIWISHSIQTG	-
N23-122	MNPNQK	IITIGSICMVVGIISLIALLN	DKHSRG	+
N17-23	MNPNQK	IITIGSICMVLQIG	NIISIWISHSIQTG	+
N17-26	MNPNQK	IITIGSICMVANSG	NIISIWISHSIQTG	-
d1 I**	MNPNQK	IITIGSVSLP	KLVEYRNWS	-
N7***	MNPNQK	IITI R SICMVVGIISLILQIG	NIISIWISHSIQTG	+
N11***	MNPNQK	IITIGSICMV D GIISLILQIG	NIISIWISHSIQTG	+
N29***	MNPNQK	IITIGSICMVVGIISLILQ R	GNIISIWISHSIQTG	+
N7-29***	MNPNQK	IITI R SICMVVGIISLILQ R	GNIISIWISHSIQTG	+
N11-29***	MNPNQK	IITIGSICMV D GIISLILQ R	GNIISIWISHSIQTG	-

FIG. 6. Summary of NA signal-anchor mutations. Shown are the amino acid sequences of the various NA mutants and their retention (+) or loss (−) of signal function. The residues in boldface represent those predicted to be involved in the hydrophobic core. *, Data taken from Davis et al. (9); **, data taken from Markoff et al. (22); ***, data taken from Sivasubramanian and Nayak (27). N11-29 was also confirmed to be negative for translocation by an in vitro transcription-translation system (unpublished data).

sequence (33), are outside the membrane-spanning region (32) of wt NA. On the other hand, in mutants like N15 and N18 these residues are able to function cooperatively with other hydrophobic sequences to span the lipid bilayer.

Analysis of a large number of signal sequences, both cleaved and uncleaved, suggests that, whereas there is no sequence specificity, there are two structural requirements for signal function. These are (i) a hydrophobic core consisting of a minimum stretch of eight apolar amino acid residues (31) and (ii) an α-helix conformation of the core (4, 11, 16). The ability of anchoring sequences (23, 41), as well as an artificially created hydrophobic peptide (40), to function as signals also supports the lack of a specific sequence requirement. In addition, flanking sequences appear to affect the translocation efficiency of the signal peptide, possibly by altering the conformation (4).

With respect to translocation, the NA mutants essentially fall into the following three classes: wt (N18 and N23-122), slightly defective (N15 and N17-23), and entirely defective (N23, N26, and N17-26). Considering the results of Spiess and Handschin (28) and Zerial et al. (40, 41), one might expect that eight or more residues of the uncharged region of NA (residues 7 to 35) would serve as a signal sequence. Therefore, it was surprising that neither N23 nor N26 retained competent signals, since they contain, respectively, 12 and 9 amino acids from this region. This behavior of N23 and N26 cannot be accounted for by the extraneous NH_2-terminal sequences, since the added residues in N23 (MPAGR) are very similar to those of N18, a translocation-positive mutant, and to those of the wt sequence (MNPNQK) in terms of overall length, charge, and the presence of a helix-breaking proline at the minus three position with respect to the charge. Similarly, N26, a translocation-negative mutant, has the identical residues present in N10, a mutant which was indistinguishable from the wt when expressed in eucaryotic cells (9). In addition, N11-29 (27), which retains amino acids 27 to 35 of wt NA, does not translocate in vitro (data not shown). Conversely, N23-122, which does not contain these residues, is as efficiently translocated as the wt protein. This, coupled with the data from the carbonate extractions, leads us to suggest that the most probable hydrophobic core sequence of NA is represented by residues 7 to 27 (32). However, a cooperative function of residues 28 to 35 with other adjacent hydrophobic sequences cannot be ruled out. On the basis of this hypothesis, we have summarized the known mutants of NA in Fig. 6. However, caution should be exercised in drawing this conclusion about the hydrophobic core region of NA, since extraneous sequences were introduced in some of these constructions. Clearly, more precise mutations are needed to test this hypothesis, and their behavior in vivo should be tested.

The data presented here indicate that the signal information is redundant within residues 7 to 27. Ten or more residues from this region appears to be sufficient for translocation, although some sequences are more efficient than others in providing the signal function. For example, a minimum of 10 residues derived from the carboxyl edge (amino acids 18 to 27) of the proposed hydrophobic core is sufficient and very efficient for translocation (N18), whereas 10 residues derived from the amino edge of this region (amino acids 7 to 16) does not supply a signal function (d1I [22] and N11-29 [27]; unpublished data derived by in vitro translocation). However, the same amino-terminal 10 residues can provide the signal function when joined to additional amino acids from the carboxyl-terminal edge (N17-23). The data also show that, although positive for translocation, N17-23 is both poorly translocated and poorly anchored, possibly indicating the loss of some important feature. This becomes more evident in N17-26, which completely loses the signal function and yet is similar to N17-23, except for residues LQI in N17-23 and ANS in N17-26. We assume that LQ to AN are conservative changes and that I to S decreases the overall hydrophobicity of N17-26. Therefore, either this slight reduction in hydrophobicity disrupts the marginal signal function of N17-23 (the residues removed in N17-23 are the most hydrophobic in the wt NA) or the ANS sequence is incompatible for the signal function. These considerations, coupled with the behavior of the point mutants we previously described (27), lead us to speculate that the helix-breaking–polar residues GS at positions 11 and 12

of the wt NA may place a structural constraint against the signal function of this region. Whereas our data are consistent with this prediction, since no mutant which included these residues as part of a minimal hydrophobic core retained a functional signal sequence, precise mutations of the GS residues at positions 11 and 12 are required to define their role and definitively establish the existence of a signal subdomain within this region. Finally, identification of a signal subdomain within a type II transmembrane domain will require evidence of a secreted protein containing an uncleaved signal in vivo.

In summary, the data presented here show that significant portions of the NH_2-terminal hydrophobic sequence are dispensable for both signal and anchor functions and suggest that the hydrophobic core is likely to be represented by residues 7 to 27. The data also show that within this region some sequences are more efficient than others in providing the translocation signal and that the wt and mutant NA proteins are translocated in the proper orientation, regardless of the characteristics of the flanking sequences.

ACKNOWLEDGMENTS

We are gratefully indebted to P. Walter and W. Hansen for the generous gift of RM and the SRP. We also thank S. Bhutani and H. Hu for assistance.

This work was supported by Public Health Service grant R01 AI16348 and postdoctoral fellowships 1 F32 GM1178801 (B.G.H.) and 5T32-CA09056 (D.J.B.) from the National Institutes of Health.

LITERATURE CITED

1. **Adams, G. A., and J. K. Rose.** 1985. Structural requirements of a membrane-spanning domain for protein anchoring and cell surface transport. Cell **41**:1007–1015.
2. **Blackburn, P., G. Wilson, and S. Moore.** 1977. Ribonuclease inhibitor from human placenta. J. Biol. Chem. **252**:5904–5910.
3. **Bos, T., A. R. Davis, and D. P. Nayak.** 1984. NH_2-terminal hydrophobic region of influenza neuraminidase provides the signal function in translocation. Proc. Natl. Acad. Sci. USA **81**: 2327–2331.
4. **Briggs, M. S., and L. M. Gierasch.** 1986. Molecular mechanisms of protein secretion: the role of the signal sequence. Adv. Protein Chem. **38**:109–180.
5. **Chou, P. Y., and G. D. Fasman.** 1978. Empirical predictions of protein conformation. Annu. Rev. Biochem. **47**:251–276.
6. **Colman, P. M., and C. W. Ward.** 1985. Structure and diversity of influenza neuraminidase. Current Top. Microbiol. Immunol. **114**:177–255.
7. **Culter, D. F., P. Melancon, and H. Garoff.** 1986. Mutants of the membrane binding domain of Semliki Forest virus E2 protein. II. Topology and membrane binding. J. Cell Biol. **102**:902–910.
8. **Dalbey, R. E., and W. Wickner.** 1988. Characterization of the signal-anchor domain of Escherichia coli leader peptidase. J. Biol. Chem. **263**:404–408.
9. **Davis, A. R., T. J. Bos, and D. P. Nayak.** 1983. Active influenza neuraminidase is expressed in monkey cells from cDNA cloned in simian virus 40 vectors. Proc. Natl. Acad. Sci. USA **80**:3976–3980.
10. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs from the VAX. Nucleic Acids Res. **12**:387–395.
11. **Emr, S. D., and T. J. Silhavy.** 1983. Importance of secondary structure in the signal sequence for protein secretion. Proc. Natl. Acad. Sci. USA **80**:4599–4603.
12. **Engelman, D. M., and T. A. Steitz.** 1981. The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. Cell **23**:411–422.
13. **Engelman, D. M., T. A. Steitz, and A. Goldman.** 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. **15**: 321–353.
14. **Fujiki, Y., L. Hubbard, S. Fowler, and P. B. Lazarow.** 1982. Isolation of intracellular membrane by means of sodium carbonate treatment: application to endoplasmic reticulum. J. Cell Biol. **93**:97–102.
15. **Hiti, A. L., and D. P. Nayak.** 1982. Complete nucleotide sequence of the neuraminidase gene of human influenza virus A/WSN/33. J. Virol. **41**:730–734.
16. **Ibrahimi, I., and R. Gentz.** 1987. A functional interaction between the signal peptide and the translation apparatus is detected by the use of a single point mutation which blocks translocation across mammalian endoplasmic reticulum. J. Biol. Chem. **262**:10189–10194.
17. **Kaiser, C. A., D. Preuss, P. Grisafi, and D. Botstein.** 1987. Many random sequences functionally replace the secretion signal sequence of yeast invertase. Science **235**:312–317.
18. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydrophobic character of a protein. J. Mol. Biol. **157**:105–132.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
20. **Lipp, J., and B. Dobberstein.** 1986. The membrane-spanning segment of invariant chain (I) contains a potentially cleavable signal sequence. Cell **46**:1103–1112.
21. **Lipp, J., and B. Dobberstein.** 1986. Signal recognition particle-dependent membrane insertion of mouse invariant chain: a membrane spanning protein with a cytoplasmically exposed amino-terminus. J. Cell Biol. **102**:2169–2175.
22. **Markoff, L., B. C. Lin, M. M. Sveda, and C. J. Lai.** 1984. Glycosylation and surface expression of the influenza virus neuraminidase requires the N-terminal hydrophobic region. Mol. Cell. Biol. **4**:8–16.
23. **Mize, N. K., D. W. Andrews, and V. R. Lingappa.** 1986. A stop transfer sequence recognizes receptors for nascent chain translocation across the endoplasmic reticulum membrane. Cell **47**: 711–719.
24. **Morimoto, T., M. Aprin, and S. Gaetini.** 1983. Use of proteases for the study of membrane insertion. Methods Enzymol. **96**: 121–149.
25. **Perara, E., and R. Lingappa.** 1985. A former amino terminal signal sequence engineered to an internal location directs translocation of both flanking protein domains. J. Cell Biol. **101**: 2292–2301.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
27. **Sivasubramanian, S., and D. P. Nayak.** 1987. Mutational analysis of the signal-anchor domain of influenza virus neuraminidase. Proc. Natl. Acad. Sci. USA **84**:1–5.
28. **Spiess, M., and C. Handschin.** 1987. Deletion analysis of the internal signal-anchor domain of the human asialoglycoprotein receptor H1. EMBO J. **6**:2683–2691.
29. **Spiess, M., and H. F. Lodish.** 1986. An internal signal sequence: the asialoglycoprotein receptor membrane anchor. Cell **44**:177–185.
30. **Tarentino, A. L., and F. Maley.** 1974. Purification and properties of an endo-β-*N*-acetylglucosaminidase from Streptomyces griseus. J. Biol. Chem. **249**:811–817.
31. **von Heijne, G.** 1985. Signal sequences: the limits of variation. J. Mol. Biol. **184**:99–105.
32. **von Heijne, G.** 1985. Structural and thermodynamic aspects of the transfer of proteins into and across membranes. Curr. Top. Membr. Transp. **24**:151–179.
33. **von Heijne, G.** 1986. Towards a comparative anatomy of N-terminal topogenic protein sequences. J. Mol. Biol. **189**:239–242.
34. **Walter, P., and G. Blobel.** 1981. Translocation of proteins across the endoplasmic reticulum, III. Signal recognition protein (SRP) causes signal sequence-dependent and site specific arrest of chain elongation that is released by microsomal membranes. J. Cell Biol. **91**:557–561.
35. **Walter, P., and G. Blobel.** 1983. Preparation of microsomal

membrane for cotranslational protein translocation. Methods Enzymol. **96**:84–93.
36. **Walter, P., and G. Blobel.** 1983. Signal recognition particle: a ribonucleoprotein required for cotranslational protein translocation. Methods Enzymol. **96**:682–691.
37. **Walter, P., and V. R. Lingappa.** 1986. Mechanism of protein translocation across the endoplasmic reticulum membrane. Annu. Rev. Cell Biol. **2**:499–516.
38. **Watson, M. E. E.** 1984. Compilation of published signal sequences. Nucleic Acids Res. **12**:5145–5164.
39. **Wickner, W. T., and H. F. Lodish.** 1985. Multiple mechanisms of protein insertion into and across membranes. Science **230**: 400–406.
40. **Zerial, M., D. Huylebroeck, and H. Garoff.** 1987. Foreign transmembrane peptides replacing the internal signal sequence of transferrin receptor allow its translocation and membrane binding. Cell **48**:147–155.
41. **Zerial, M., P. Melancon, C. Schneider, and H. Garoff.** 1986. The transmembrane segment of human transferrin receptor functions as a signal peptide. EMBO J. **5**:1543–1550.