The Role of the N Region in Signal Sequence and Signal-anchor Function*

To determine the role of sequences other than the hydrophobic core in mediating signal sequence function, we examined the behavior of fusion proteins and deletion mutants in cell-free systems. We demonstrate that neither the N nor the C region of the preprolactin signal sequence is necessary for translocation. However, insertion of sequences with either a net charge of +2.5 or -6.0 between the N region and the hydrophobic core of the signal converted it into a signal-anchor. The topologies adopted (types I and II, respectively) were opposite those predicted from the distribution of charges surrounding the hydrophobic core of the signals. When these mutant signals were located in the interior of an otherwise secreted protein, both sequences functioned as stop-transfer sequences. Related mutations were assayed in fusion proteins in which the IgM transmembrane domain functioned as an amino-terminal signal-anchor. For these molecules, the distribution of charged residues surrounding the hydrophobic core had no influence on the topology adopted. Our results suggest that features other than simple charge distribution play an important role in determining membrane topology in vitro.

The topogenic elements that mediate translocation across or integration into the ER membrane comprise three classes. First, secretory signals are generally amino-terminal extensions that are cleaved from the nascent polypeptide during or immediately after transport across the ER membrane. Second, signal-anchor sequences initiate translocation of polypeptide domains; but in contrast to secretory signal sequences, they are neither cleaved nor released into the lumen of the ER. Third, stop-transfer sequences normally follow either a signal or signal-anchor sequence and function to abort the translocation process and to integrate the polypeptide in the ER membrane. As a consequence of the topogenic properties of these sequences, the transmembrane domains of integral membrane proteins consist primarily of signal-anchor and stop-transfer sequences.

Statistical analysis of a large number of sequences revealed several characteristics shared by most eukaryotic signals (1-4). These features divide secretory signal sequences into three distinct domains termed N, H, and C regions (1). The N region is typically positively charged and located at the amino-terminal end. The H region is the hydrophobic core of the sequence and contains a stretch of 7-15 uncharged amino acids. The C region is involved in recognition and cleavage by signal peptidase. It is not known whether the N and H regions represent functionally distinct domains.

The functional properties of specific subsequences of signal peptides have been examined using mutant signals in vitro and in vivo (5-7). Analysis of such mutants has demonstrated the importance of the hydrophobic core sequence (8) and confirmed the requirement for specific amino acids in the cleavage site (9). Fusion proteins have proven very useful for characterizing both normal and mutant signal sequences (10-12). Fusion proteins have also been used to illustrate the tremendous variety of sequences that, when positioned in an appropriate context, can functionally replace the signal sequence of yeast invertase (13). More recently, measurements of relative efficiency have been used to characterize the role of core sequence hydrophobicity (14-17) and of sequences carboxyl-terminal to the cleavage site (18, 19) in translocation function.

Mutations in the C region that abolish cleavage by signal peptidase also prevent complete translocation of some (20), but not all (19), polypeptides, resulting in molecules that remain anchored in the ER membrane (20). Although some of these mutant signals function as signal-anchor sequences, lack of signal cleavage cannot be the only determinant of membrane integration as both naturally occurring and genetically manipulated signals have been characterized that are not cleaved, but are completely translocated and released into the ER lumen (17, 19, 21-23).

Signal-anchor sequences can mediate the insertion of a membrane-spanning sequence with either type I (amino-terminal extracytoplasmic) or type II (amino-terminal cytoplasmic) orientation. In comparison, stop-transfer sequences do not normally initiate the translocation of a polypeptide domain. For this reason, there is no requirement for stop-transfer sequences to specify topology. As a result, stop-transfer sequences span the membrane in either orientation as dictated by the other topogenic elements in the polypeptide.

The stop-transfer sequence of the IgM \( \mu \) heavy chain can function as a signal-anchor when expressed as the amino-terminal topogenic element of a fusion protein (24). Similar to other characterized signal-anchor sequences, when the IgM stop-transfer sequence functions as a signal-anchor, the nascent protein is targeted to the ER membrane via interactions with SRP and the SRP receptor (25). However, the fidelity with which a single membrane topology is achieved has not been carefully investigated.

* This work was supported by Medical Research Council of Canada Grant MT-10490 and a Medical Research Council of Canada scholarship (D. W. A.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed. Tel.: 416-525-9140 (ext. 2075); Fax: 416-522-9033.

‡ The abbreviations used are: ER, endoplasmic reticulum; SRP, signal recognition particle; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine.
The mechanism by which a signal-anchor sequence adopts a particular orientation is unknown. However, analysis of a large number of transmembrane domains suggests that the amino acids flanking the hydrophobic core of the most amino-terminal signal-anchor determine which membrane topology is adopted (26-30). One hypothesis for the underlying mechanism contends that negatively charged amino acids are intrinsically more easily translocated than positive ones (26).

The principal attraction of this model is that it accounts for the preponderance of positive charges observed in the cytoplasmic loops of proteins that span the membrane multiple times (31, 32). In addition, some direct biochemical evidence for this model has been presented (26). A related model suggests that it is not the absolute number of charges, but the difference in net charge between the amino terminus and carboxyl terminus that determines the orientation of a signal-anchor in the ER membrane (27). As formalized by Hartmann et al. (27), this model predicts that the orientation of the signal-anchor closest to the amino terminus of a protein is defined by the net charge in the 15 amino acids on either side of the hydrophobic core of the sequence.

Charged amino acids are also nonrandomly distributed in secretory signal sequences where the average net charge of the amino terminus is +1 and the carboxyl-terminal end is uncharged (1). The charge distribution is similar to that found in non-type II signal-anchor sequences. Therefore, mechanisms must exist that determine membrane orientation and differentiate uncleaved secretory signals from signal-anchor and stop-transfer sequences. This issue is particularly relevant for the IgM transmembrane domain because of the relatively low hydrophobicity of the core sequence of this molecule compared to many secretory signal sequences (33).

Here we examine mutations in the sequence of amino acids flanking the hydrophobic core of the preprolactin signal sequence that are sufficient to convert the signal into either a type I or II signal-anchor. However, in both cases, the topology adopted was opposite to that predicted from charge distribution. To address the issue of membrane orientation in more detail, we used fusion proteins to examine the importance of the charged regions flanking the IgM transmembrane domain for signal-anchor function. Changes in the distribution of these charged residues did not alter the orientation adopted when the sequence functioned as an amino-terminal signal-anchor. Our results suggest that factors other than simple charge distribution and hydrophobicity influence the orientation of amino-terminal signal-anchor sequences.

**EXPERIMENTAL PROCEDURES**

**General Methods**—Restriction endonucleases and other DNA-modifying enzymes were purchased from New England Biolabs, Inc. or from Boehringer Mannheim and were used according to the manufacturers’ instructions. RNase inhibitor (RNA Guard) was from United States Biochemical Corp., and sodium thiocyanate (Aldrich), and 0.2 M sodium thiouracil (ultrapure-grade, ICN Biomedicals), 8.0 M sodium thiocyanate (Aldrich), and 0.2 M sodium carbonate (ACS reagent-grade, Sigma) were prepared immediately before use.

**Recombinant DNA Constructs**—Construction of the plasmid encoding the parent molecule, composed of 117 residues of chimpanzee α-globin, containing a glycosylation site fused to the amino terminus of the complete coding region of bovine preprolactin or the Pt passenger (amino acids 58-199 of prolactin) has been described (18, 24). Deletion of the C region of the preprolactin signal sequence and construction of the plasmid encoding the deleted signal sequence fused to the Pt passenger have also been described (18).

Construction of the plasmid encoding the complete IgM transmembrane domain fused to the amino terminus of the Pt passenger (referred to here as XA-HiKPt) has also been described (24). Initial characterization of the signal-anchor function of the IgM stop-transfer sequence in this context was described (18). Further manipulations of these plasmids are described below.

**Mutant Signal Sequences**—To facilitate the construction of insertion mutations in the NHp signal, two single nucleotide changes were introduced in a plasmid encoding a deletion mutant of the preprolactin signal sequence fused to the Pt reporter domain (previously described as pSPs+PtP (18)) using the methods described by Gunluk et al. (36). To permit insertion of the A- region, the sequence of the AuaII restriction site in pSPs-PtP (GGTCC) was changed to GGGAC to match the naturally occurring AuaII sites at each end of the A- coding region. This modification does not alter the amino acid sequence encoded by the plasmid. The second mutation was a single base substitution that results in the replacement of the Arg at the amino-terminal end of the hydrophobic core of the preprolactin signal with Cys. This mutation removes one positive charge from the signal and generates the plasmid pSPNHpPc, which encodes the NHp signal (described in Fig. 1) fused to the Pt passenger domain. The A- region was then excised as an AuaII fragment from a plasmid encoding the IgM stop-transfer sequence and ligated into the modified AuaII site of pSPNHp. The A+ region results from inserting the DNA encoding the A- region into the plasmid encoding the unmodified AuaII site in the NHp signal. Because the central residue of the AuaII sites in the two plasmids is not the same, the fragment is inserted with the opposite orientation. The amino acids encoded by the fragment in this orientation are shown in Fig. 1. DNA encoding each of the mutant signals was then inserted into pSPs-PtP in place of the existing signal sequence.

To add the wild-type preprolactin signal to the gG passenger domain, we employed an intermediate used previously to generate carboxyl-terminal signal deletions (18). This intermediate has a unique AccI site in the plasmid at the carboxyl-terminal end of the region encoding the preprolactin signal sequence. After digesting the plasmid with AccI, the ends of the DNA were blunted with mungbean nuclease and ligated to a similarly blunted NcoI site at the amino terminus of the domain encoding gG to make the plasmid pSpPtP462. This deletes the amino-terminal methionine from the gG domain and thereby reduces the frequency of internal initiation of translation when the plasmid is transcribed and translated in vitro. The DNA encoding SpgG was then added to the amino terminus of the coding regions for the mutant signals fused to the Pt reporter. The SpgG coding region was digested with BstEII, end-repaired with the Klenow fragment of DNA polymerase I, and subsequently digested with NheI. This fragment was ligated into the plasmids encoding the mutant signal sequences with NheI and restriction enzyme recognition sites: an NcoI site encoding the amino-terminal signal-anchor closest to the amino terminus of a protein is is adopted.

**Signal-anchor Sequences**—The DNA encoding the IgM transmembrane domain fused to the Pt passenger includes several useful restriction enzyme recognition sites: an NcoI site encoding the amino-terminal signal-anchor, two AccI sites flanking the coding sequence for the A- region (Fig. 1), an EcoNI site just 5′ of the sequence encoding the K region (Fig. 1), and an EcoRI site at the 3′ end of the stop-transfer coding region.

To facilitate use of the AuaII sites, the coding region for XA-HiKPt was subcloned into a vector without endogenous AuaII sites. The resulting plasmid was named pKPXA-HiKPt. There are two AuaII sites in the coding sequence that confers resistance to ampicillin (β-lactamase); therefore, the plasmid employs resistance to kanamycin for selecting transformants. To construct XHHiKPt, the sequence encoding the A- region was removed from the plasmid pKPXA-HiKPt by digestion with AuaII. To construct NHiKPt, the plasmid pKPXA-HiKPt was digested with NcoI and AuaII to remove the sequences encoding the X and A- regions. These sequences were then filled in with the coding region for N (with the restriction enzyme recognition site) from the plasmid encoding the NHp signal sequence described above. To delete the K region from the IgM stop-transfer sequence, two partially complementary oligonucleotides were synthesized as an adaptor and inserted into the plasmid encoding XHHiKPt and the 5′-end of the coding region for N. The K region was then assembled into the pEcoRI site with the frame adjusted to match the coding region of the EcoRI site at the 3′-end of the K coding region. Therefore, digestion of this plasmid with EcoRI deleted the K coding region precisely. The coding regions...
for each of the mutant signal-anchors fused to Pt were then cloned (as NcoI/PstI fragments) into the plasmids encoding the wild-type proprolactin signal fused to the gG and Pt coding domains.

All of the resulting plasmids were mapped extensively with restriction endonucleases, and most were sequenced from the 5’-end into the region encoding Pt before use. Complete details of any of the plasmids are available from us.

Transcription-linked Translation—Transcription of SP6 plasmids was as described previously (37). Aliquots of the transcription reaction mixture were added directly to translation reactions at a final concentration of 5% as described (37). Proteins synthesized in vitro were labeled by including [35S]methionine in the transcription reaction and visualized by fluorography after separation by SDS-PAGE using either the standard Laemmli (38) or the Tris/Tricine (39) buffer system.

Orientation of Membrane Proteins—Protease protection and sedimentation assays were performed as described previously (40). Solubilization of the membranes in the absence of SRP/SRP receptor targeting and translocation system.

Translation reactions for preprolactin and XA-HiKPt were included as secretary and integral membrane controls, respectively, in every assay of membrane integration. To control for nonspecific aggregation of the fusion proteins in the high pH extraction buffers, some samples were processed without added microsomes. Other controls designed to demonstrate that membrane association required the sequences inserted between the N and Hp regions. The microsomes were solubilized directly in SDS-PAGE loading buffer and incubated at 65 °C for 5 min. Bovine serum albumin (1.0 μg) was added to the supernatant fractions (170 μl) as carrier, and proteins were collected by precipitation with ice-cold trichloroacetic acid (final concentration of 15%) and prepared for SDS-PAGE as above.

Translation reactions for preprolactin and XA-HiKPt were included as secretory and integral membrane controls, respectively, in every assay of membrane integration. To control for nonspecific aggregation of the fusion proteins in the high pH extraction buffers, some samples were processed without added microsomes. Other controls designed to demonstrate that membrane association required the sequences inserted between the N and Hp regions. The microsomes were solubilized directly in SDS-PAGE loading buffer and incubated at 65 °C for 5 min. Bovine serum albumin (1.0 μg) was added to the supernatant fractions (170 μl) as carrier, and proteins were collected by precipitation with ice-cold trichloroacetic acid (final concentration of 15%) and prepared for SDS-PAGE as above.

Translation reactions for preprolactin and XA-HiKPt were included as secretory and integral membrane controls, respectively, in every assay of membrane integration. To control for nonspecific aggregation of the fusion proteins in the high pH extraction buffers, some samples were processed without added microsomes. Other controls designed to demonstrate that membrane association required the sequences inserted between the N and Hp regions. The microsomes were solubilized directly in SDS-PAGE loading buffer and incubated at 65 °C for 5 min. Bovine serum albumin (1.0 μg) was added to the supernatant fractions (170 μl) as carrier, and proteins were collected by precipitation with ice-cold trichloroacetic acid (final concentration of 15%) and prepared for SDS-PAGE as above.
NA+HpPt molecules were translocated and therefore protected from protease (lane 11). In contrast to these molecules, when microsomes were added to translocation reactions synthesizing NA−HpPt, signal cleavage was not observed (lane 14). Instead, the band corresponding to NA−HpPt shifted to lower molecular weight only after protease was added to the reaction, suggesting that the molecules adopted a transmembrane orientation (lane 15). This behavior is similar to that obtained when the stop-transfer sequence from the IgM heavy chain (XA−HiK) was fused to the same passenger (lanes 17–20; see also Ref. 24). Therefore, these results suggest that addition of the A− region to the NHp secretory signal sequence converted it to a type II signal-anchor (compare lanes 13–16 with 17–20).

The topology observed for both NA−HpPt and XA−HiKPt (type II, amino-terminal cytoplasmic) is opposite that predicted from the distribution of charged residues in the amino acid sequence of the molecules. To examine this issue in greater detail, membrane topology was assessed for three additional fusion proteins, XHiKPt, NHiKPt, and XHiPt. These molecules differ only in those residues adjacent to the hydrophobic core of the IgM stop-transfer sequence. In all of these molecules, the highly negatively charged A− region has been deleted. The amino terminus of NHiKPt was derived from the preprolactin signal (Fig. 1) and is positively charged. In XHiPt, the endogenous charge clusters have been deleted from both ends of the IgM stop-transfer sequence.

Standard proteolysis reactions demonstrated that even though the distribution of charges varies widely, all of the mutant signal-anchors translocated the Pt domain across the ER membrane (Fig. 3). With this assay, it is not possible to determine whether or not these molecules remain transmembrane as they are not exposed to the protease (compare lanes 2 and 3 with 6 and 7 and 11 and 12). However, we demonstrate below that these molecules are tightly associated with the ER membrane. Translocation of the Pt domain is slightly more efficient for NHiKPt than for the other molecules. Although the difference is small, the result is reproducible. Therefore, it is possible that the N region from the preprolactin signal sequence increases recognition of the topogenic element.

Because targeting to the membrane was <100% efficient for both the signal sequence mutants (NA−HpPt, NA+HpPt, and NA+A+HpPt) described above and the signal-anchor mutants (XHiKPt, NHiKPt, and XHiPt), it is possible that some molecules adopted both type I and II transmembrane topologies. Molecules with type I orientation would remain accessible to protease and therefore cannot be differentiated from nontargeted molecules in this assay.

To determine membrane orientation unambiguously, molecules that had functionally interacted with the membranes were collected from the reticulocyte lysate translation reactions by centrifugation, resuspended, and subjected to proteolysis. Orientation was determined by comparing the percent protection measured for the mutants with that of the control signal-anchors (lanes 4 and 5). In contrast, NA+HpPt molecules (data not shown) adopt two distinct orientations (lanes 7–9). Those molecules cleaved by signal peptidase (lane 7, lower band) were located in the lumen of the ER and were therefore protected from protease. However, those molecules that were not cleaved (lane 7, upper band) remained sensitive to protease, consistent with a type I orientation.
The Role of the N Region in Signal Sequence Function

In contrast, ~30% of each of the fusion proteins bearing the mutant signal-anchors (Fig. 4B, XA–HiKPt (lanes 1–3), NHiKPt (lanes 4–6), and XHiPt (lanes 7–9)) was protected from added protease, suggesting that molecules containing the Hi transmembrane sequence as a signal-anchor adopt exclusively the type II topology. The topologies assigned to the signal and signal-anchor mutants were confirmed using an alternate approach in which microsomes were collected by centrifugation in parallel with proteolysis of complete translation reactions (data not shown). Moreover, when the amount of microsomes added to translation reactions was doubled to substantially increase targeting efficiency (at the expense of translational efficiency), proteolysis assays confirmed that all three signal-anchor mutants adopt primarily the type II orientation (data not shown).

The relatively low hydrophobicity of the IgM transmembrane domain suggested that some of the mutant signal-anchor sequences may function as uncleaved signal sequences. In the absence of signal cleavage, protease protection assays do not permit molecules that have been fully translocated and released into the ER lumen to be distinguished from those that remain integrated in the ER membrane with the Pt domain located on the luminal side. Moreover, protease protection assays do not distinguish peripheral from integral membrane proteins. Therefore, it is also possible that some of the fusion proteins were translocated, but remained lodged in the translocation machinery rather than integrated in the ER membrane. To distinguish these possibilities, microsomes were collected from the translation reactions and incubated in 200 mM Tris acetate (pH 9.0) containing 0.08% sodium deoxycholate. After 30 min at 0 °C, the insoluble material was collected by centrifugation. Control experiments demonstrated that this treatment was sufficient to release peripheral membrane proteins as well as the soluble contents of the ER lumen (data not shown). Hence, if the protease-sensitive uncleaved NA+HpPt and NA+A+HpPt molecules result from an aborted translocation event, these molecules should be sensitive to extraction. Similarly, if XHiKPt, NHiKPt, and XHiPt are protected from protease because they are fully translocated and released in the ER lumen, they will be recovered in the supernatant fractions using this assay.

After extraction, most of the XA–HiKPt and NA–HpPt molecules were found in the membrane pellet (76%), consistent with our interpretation that these molecules fully integrated in the ER membrane (Fig. 5, lanes 1–4). In contrast, those molecules of NA+HpPt and NA+A+HpPt completely translocated and cleaved by signal peptidase (lanes 5–8, lower bands) were largely released from the microsomes by the extraction (~68%). However, the nontranslocated (protease-sensitive) molecules (lanes 5–8, upper bands) remained associated with the membranes (~72%), similar to the transmembrane molecule XA–HiKPt. In addition, XHiPt (lanes 9 and 10) as well as NHiKPt and XHiKPt (data not shown) were observed primarily in the membrane pellet, confirming that these molecules are also tightly associated with the ER membrane. Similar results were obtained when microsomes were extracted with sodium carbonate (pH 11.5) in combination with either urea or sodium thiocyanate (data not shown). Moreover, in control experiments (see “Experimental Procedures”) in which microsomes were added after translation was terminated (or employing microsomes inactivated for SRP-dependent translocation), none of the molecules bound tightly to microsomes (data not shown). Together, these results strongly suggest that the uncleaved NA+HpPt and NA+A+HpPt molecules are integrated in the ER membrane with type I topology, whereas XA–HiKPt, XHiKPt, NHiKPt, and XHiPt adopt the type II integral membrane topology.

The charge distribution calculated for the 15 amino acids on either side of the hydrophobic core (as suggested by Hartmann et al. (27)) is presented along with the predicted and
The Role of the N Region in Signal Sequence Function

observed topologies of each of these mutants in Table I. It is obvious that the charge distribution calculated in this manner cannot be used to predict the topologies of these mutants. In addition to the data presented in Table I, there are several characteristics of the fusion proteins that are potentially relevant to the topology adopted. The amino terminus of the wild-type IgM transmembrane domain is longer than 15 amino acids, contains several other charged residues, and has a net charge of −5.0. Deletion of the A− segment removes the cluster of Gln residues responsible for most of the negative charges in XA−HiKPT. The complete amino-terminal end of XA−HiKPT was replaced with the amino terminus of the preprolactin signal sequence in NHiKPT. In addition to effectively reversing the charge distribution of the amino terminus, this region may include a feature involved in some aspect of signal sequence recognition. Finally, the endogenous charge clusters from both ends of the IgM transmembrane domain were deleted to generate the fusion protein XHiPt. Deleting the charged residues from the carboxyl-terminal end moves four negative charges and one positive charge within 15 residues of the hydrophobic core of the transmembrane domain that are just beyond this region in the other mutants.

Many signal-anchor sequences function as stop-transfer sequences when presented as part of a translocating peptide. To assay whether any of the mutant signal sequences would function as a stop-transfer sequence, a series of fusion proteins was constructed with the mutant signals positioned between the globin and Pt passenger domains. Translocation of these molecules was initiated by a wild-type preprolactin signal located at the amino terminus of the globin domain. To provide an additional assay for translocation, the globin domain included an 8-amino acid glycosylation site inserted at residue 19. If the mutant signals functioned as stop-transfer sequences, then the molecules would anchor in the ER membrane with the Pt domain on the cytoplasmic side, where it would be accessible to added protease. However, the translocated globin domain would be glycosylated and protected from protease.

Most of the SpgGNHpt molecules with the internalized NHp signal were fully translocated and protected from protease (Fig. 6A, lanes 1–5). When microsomes were added to the reactions (lanes 3 and 4), the two major bands observed resulted from cleavage of the amino-terminal signal and differential glycosylation of the translocated molecules. The upper and lower bands represent glycosylated (downward pointing arrowhead) and nonglycosylated (upward pointing arrowhead) molecules, respectively (lane 3). Addition of proteinase K revealed that almost 80% of the SpgGNHpt molecules were fully translocated without cleavage of the internalized NHp signal. Densitometry (lanes 4 and 6) suggests that ~10% of the protease-protected prolactin immunoreactive molecules were released from SpgGNHpt by signal peptidase cleavage (lane 6, dot). Similar analysis of the bands corresponding to the glycosylated globin domain indicated that ~20% of the globin domains were not associated with prolactin after proteolysis (lane 5, open arrowhead). Therefore, no more than 10% of the SpgGNHpt molecules adopted a transmembrane topology. These results demonstrate that when the NHp signal was not cleaved from the translocating polypeptide, it did not abort translocation or mediate membrane integration.

In contrast, ~90% of the translocated SpgGNA−HpPt molecules (Fig. 6B) and 70% of the translocated SpgGNA+Hpt molecules (Fig. 6C) adopted a transmembrane topology. As expected, proteolysis released both glycosylated and nonglycosylated globin domains (Figs. 6, B and C, lanes 5, open and closed arrowheads, respectively). Moreover, cleavage was not observed for either the internalized NA−Hpt or NA+Hpt topogenic element as the prolactin-reactive species present after proteolysis co-migrated with the nonproteolysed molecules (Fig. 6, B and C, compare lanes 4 and 6). Although visual inspection might suggest that a smaller fraction of SpgGNA+Hpt molecules adopt a transmembrane topology, the number of Met residues in full-length molecules (6) compared to those in globin domains released by proteolysis (2) accounts for the relative density of the bands on the film. Parallel proteolysis assays for similar molecules with either the XA−HiK or XHi element in place of the internalized signal resulted in ~10% residual full-length molecules after translocation and proteolysis (data not shown). Therefore, we conclude that the NA−Hpt element functions to halt translocation as efficiently as XA−HiK (the wild-type IgM stop-transfer sequence). Moreover, the NA+Hpt element can also function as a stop-transfer element, albeit with reduced efficiency.

To determine whether the NA−Hpt element integrates the nascent SpgGNA−HpPt polypeptide stably in the ER membrane, microsomes were extracted with chaotropic agents as described above for NA−Hpt. Extraction with 100 mM sodium carbonate (pH 11.5) combined with 1 M Urea (Fig. 7A, lanes 1–6) or 1 M NaSCN (lanes 7–12) as well as 200 mM Tris (pH 9.0), 0.08% deoxycholate (Fig. 7B, lanes 1–6) demonstrated that molecules anchored via the NA−Hpt sequence were as tightly associated with microsomes as were those anchored by the wild-type IgM stop-transfer sequence (XA−HiK). In contrast, all of these extraction conditions released a large fraction of the fully translocated SpgGNA+Hpt molecules from the lumen of microsomes. Control experiments in which microsomes were added after translation had terminated or in which microsomes were added that had been inactivated for translocation (as described above) demonstrated that membrane association depends on functional translocation machinery in the ER membrane.

Discussion

Mutant Signal Sequences—Deletion of the C region from the preprolactin signal sequence had no effect on translocation or signal cleavage for the molecule NHppt, yet the deletion abolished cleavage when this signal was positioned internally in SpgGNHpt. Nevertheless, SpgGNHpt was fully translocated and released into the ER lumen (Figs. 5A and 6). This result demonstrates that lack of cleavage of the internalized signal sequence is not sufficient to account for membrane integration via the NHp topogenic element. Therefore, it is reasonable to expect that loss of signal cleavage is also not sufficient to cause membrane integration of those molecules with a modified NHp signal at the amino terminus (NA−Hpt and NA+Hpt). Moreover, it strongly suggests

| Table I | Predicted topologies of fusion proteins |
|---------|---------------------------------------|
| Fusion protein | ∆N | ∆C | ∆(C − N) | Topology | Predicted | Observed |
| XA−HiKPT | −5.0 | +1.5 | +6.5 | I | II |
| XHiKPT | +1.0 | +1.5 | +0.5 | I | II |
| NHiKPT | +2.0 | +1.5 | −0.5 | II | II |
| NHiKPT | +1.0 | −2.5 | −3.5 | II | II |
| NA+A+Hpt | +1.5 | −1.5 | −3.0 | II | II |
| NA+Hpt | +1.5 | −1.5 | −3.0 | II | II |
| NA−Hpt | −5.0 | +1.5 | +3.5 | I | II |

*Approximately 58% of the Na+Hpt molecules and 48% of the NA+A+Hpt molecules were signal-cleaved and secreted.
that hydrophobicity is not the only determinant of stop-transfer activity.

Deletion of the N region had no effect on translocation for the molecule HPrPt, suggesting that this region does not contain information specifically required to initiate translocation or for recognition by signal peptidase. However, insertion of sequences between the N and Hp subdomains of the NHp signal sequence had profound effects on translocation, signal cleavage, and membrane topology.

Addition of the A- subdomain to the NHp signal converted this normally secreted signal into a type I signal-anchor (Figs. 2, 4, and 5). The A- region was derived from the IgM stop-transfer sequence, where it is located immediately amino-terminal to the hydrophobic transmembrane domain. The most striking feature of this sequence is the large net negative charge (-6.0). Current models for membrane topology predict (26-30) that addition of negative charges to the amino terminus of a signal-anchor sequence would result in type I topology, precisely opposite the topology adopted by NA-HpPt (Fig. 4). Although it is possible that this molecule adopts the type II topology because the A- segment is inherently difficult to translocate, such a result would be surprising given that this segment is normally translocated in IgM and is also translocated across the ER membrane in the molecule SpgGNA-HpPt.

The A- segment contains 21 amino acids; therefore, it is possible that some feature other than the 6 Glu residues contributes to the transmembrane topology of the mutant signals examined here. For example, the A- region adds Leu and Trp to the amino-terminal end of the hydrophobic core of the signal. These residues are conserved at a similar location in the transmembrane region of several classes of immunoglobulins (42-46). Moreover, adding these residues also increases the hydrophobicity of the NHp element. However, the NHp signal contains a 5-amino acid deletion and therefore has lower average and total hydrophobicities than the wild-type preprolactin signal. Calculations based on several different scales of hydrophobicity (48) indicate that adding the Leu and Trp residues does not increase the hydrophobicity of NHp enough to offset the 5-amino acid deletion. Furthermore, dramatically increasing the hydrophobicity of a signal sequence does not necessarily impair secretion (14-16). Therefore, it is unlikely that hydrophobicity contributes to the observed phenotype. However, it is possible that the reduced hydrophobicity of NA+Hp (due to the Cys to Arg mutation) contributes to the fraction of molecules that are fully translocated and released from the ER membrane. In addition, it is not known whether the lack of signal cleavage associated with these mutants results from altered recognition of signal peptidase or is a consequence of the altered topology. Experiments to delineate these issues further are in progress.

Both NA+HpPt and NA+A+HpPt have a net positive charge at the amino terminus and are therefore predicted to
orient with the type II topology. Nevertheless, the fraction of these molecules that are not fully translocated and released into the ER lumen (42% for NA+HpPt and 52% for NA+A+HpPt) appear to be integrated in the ER membrane with type I topology (Fig. 4). Similar to the A− segment, the A+ sequence contains amino acids (6 prolines) relatively uncommon in the N region of signal sequences (4, 49). Nevertheless, our results demonstrate that these residues are not inherently incompatible with either cellular targeting (via SRP/SRP receptor) or translocation machinery. Moreover, duplication of the inserted sequence yielded an efficiently targeted molecule containing 12 prolines in the N region. It is difficult to rationalize the topology of NA+HpPt and NA+A+HpPt resulting from the inability to translocate the A+ domain at the amino terminus of the fusion protein because these molecules adopt the type I orientation. Although we cannot determine the location of the proline-rich amino terminus of the molecule experimentally, it is unlikely that it remains on the cytoplastic side of the ER membrane because the 13-residue hydrophobic segment (Hp) is not long enough to span the membrane twice. Moreover, molecules with one or more copies of the A+ segment in the NHp signal contain two positive charges immediately adjacent to the amino-terminal end of the hydrophobic Hp segment. Therefore, even if the hydrophobic segment is extended to include the first 8 amino acids of the Pt passenger (including 1 His, 2 Thr, and 2 Ser residues; see Fig. 1), the maximum length of the potential transmembrane region is only 21 amino acids. Nevertheless, it remains to be demonstrated whether the integral membrane topology of NA+HpPt and NA+A+HpPt is governed by the distribution of charges, prolines, or some other feature of the sequence.

Duplicating the A+ segment increased the targeting efficiency as well as the relative proportion of molecules that adopted type I membrane orientation. Nevertheless, 48% of the NA+A+HpPt molecules are cleaved by signal peptidase and released into the ER lumen. These results corroborate earlier observations that a variety of segments including sequences that diverge widely from the consensus characteristics for a signal sequence can still function in translocation (13).

**Mutant Signal-anchor Sequences**—Neither the orientation nor the membrane integration function of the IgM transmembrane domain was affected by the sequence of amino acids flanking the hydrophobic core sequence. Therefore, membrane integration with type II orientation requires only features within the HI segment (the hydrophobic core of the IgM transmembrane domain). The most obvious mechanism for interaction of the HI segment with the ER membrane is one based on hydrophobicity. However, both the average and total hydrophobicities of the HI segment (1.6 and 37, respectively; calculated according to Ref. 50) are less than those of many signal sequences (1). For example, the corresponding values for the preprolactin signal sequence are 2.4 and 41.4, respectively.

One possible explanation for the observation that only type II topology is observed when the HI segment functions as a signal-anchor is that integration in this orientation may occur in the absence of signals specifying type I topology. Hartmann et al. (27) have also suggested a default pathway leading to type II orientation as a possible explanation for molecules that do not integrate with the orientation predicted from the distribution of charges surrounding the amino-terminal signal-anchor. Such a model does not preclude the existence of signals specifying type II topology for some molecules. If this model is correct, our data predict that a signal involved in mediating type I topology occurs (fortuitously) in the mutant signals that contain one or two copies of the A+ segment. Experiments to test this possibility are in progress.

Although our results suggest that the signal for orientation is not simple charge distribution, they do not rule out the possibility that charged residues contribute to the features that specify topology. To the best of our knowledge, NA−Hp is the only example of a secretory signal that has been completely converted to a signal-anchor by a mutation in the N region that does not alter the hydrophobicity of the core sequence. Alterations in the region amino-terminal to the hydrophobic core of other topogenic elements have been shown to result in polypeptides that adopt more than one topology (6, 12, 30, 50−52). For a few of these, both secreted and integral membrane forms were obtained, similar to the results shown here for NA+HpPt and NA+A+HpPt (6, 50). Although phenotypically similar to the signals with one or two copies of the A+ domain, comparison of the sequences of all of these topogenic elements did not reveal any obvious similarities. In at least two cases, addition of a segment with a net negative charge to the amino terminus of a type II signal-anchor resulted in a molecule that adopted both type I and II integral membrane topologies (30, 52). These results, combined with the results presented here, are difficult to reconcile with any mechanism for directing membrane orientation based on charge distribution alone.

**Acknowledgments**—We are grateful to Alison Cowie and Fabiola Janiak for critical reading of the manuscript and to Domina Falcone for technical assistance.

REFERENCES

1. Von Heijne, G. (1985) J. Mol. Biol. 184, 99−105
2. Shinde, U. P., Guru Row, T. N., and Mawal, Y. R. (1989) Biosei. Rep. 9, 737−745
3. Prabhakaran, M. (1990) Biochem. J. 269, 691−696
4. Landry, S. J., and Gierasch, L. M. (1991) Trends Biochem. Sci. 16, 159−163
5. Ibrahimi, I., and Gentz, R. (1987) J. Biol. Chem. 262, 10189−10194
6. Szczezna-Skorupa, E., Mead, D. A., and Kemper, B. (1987) J. Biol. Chem. 262, 8966−8970
7. Green, R., Kramer, R. A., and Shields, D. (1989) J. Biol. Chem. 264, 2969−2974
8. Bird, P., Gething, M. J., and Sambrook, J. (1990) J. Biol. Chem. 265, 8420−8425
9. Sturzaker, S. C., Whittfeld, P. L., Christie, D. L., Bellamy, R. A., and Both, G. W. (1987) J. Cell Biol. 105, 2897−2903
10. Lingappa, V. R., Chaidez, J., Yost, C. S., and Hedgrehp, J. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 456−460
11. Lipp, J., and Dobberstein, B. (1988) J. Cell Biol. 106, 1813−1832
12. Yost, C. S., Lopez, C. D., Prusiner, S. B., Myers, R. M., and Lingappa, V. R. (1980) Nature 284, 699−701
13. Kaiser, C. A., Preuss, D., Grisafi, P., and Botstein, D. (1987) Science 235, 312−317
14. Chou, M. Y., and Kendall, D. A. (1990) J. Biol. Chem. 265, 2873−2880
15. Kendall, D. A., Doud, S. K., and Kaiser, E. T. (1990) Biopolymers 29, 139−147
16. Cloifi, J. A., Allen, K. L., Lively, M. O., and Kemper, B. (1989) J. Biol. Chem. 264, 15052−15058
17. von Heijne, G., Liljestrom, P., Mikus, P., Anderson, H., and Ny, T. (1991) J. Biol. Chem. 266, 15240−15243
18. Andrews, D. W., Perrera, E., Lesser, I., and Lingappa, V. R. (1988) J. Biol. Chem. 263, 17091−17098
19. Kohara, A., Yamamoto, Y., and Kikuchi, M. (1991) J. Biol. Chem. 266, 20363−20368
20. Shaw, A. S., Rottier, P. J. M., and Rose, J. K. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 7592−7596
21. Bonatti, S., and Blobel, G. (1979) J. Biol. Chem. 254, 12261−12264
22. Lingappa, V. R., Lingappa, J. R., and Blobel, G. (1979) Nature 281, 711−719
The Role of the N Region in Signal Sequence Function

23. Eble, B., MacRae, D. R., Lingappa, V. R., and Ganem, D. (1987) Mol. Cell. Biol. 7, 3591–3599
24. Rothman, R. E., Andrews, D. W., Calayag, M. C., and Lingappa, V. R. (1988) J. Biol. Chem. 263, 10470–10480
25. Mize, N. K., Andrews, D. W., and Lingappa, V. R. (1986) Cell 47, 711–719
26. Von Heijne, G. (1989) Nature 341, 456–458
27. Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 5786–5790
28. Dalbey, R. E. (1990) Trends Biochem. Sci. 15, 253–257
29. Parks, G. D., and Lamb, R. A. (1991) Cell 64, 777–787
30. Beltzer, J. P., Fiedler, K., Fuhrer, C., Geffen, I., Handschin, C., Wessels, H. P., and Spiess, M. (1991) J. Biol. Chem. 266, 973–978
31. von Heijne, G. (1986) EMBO J. 5, 3021–3027
32. von Heijne, G., and Gavel, Y. (1988) Eur. J. Biochem. 174, 671–678
33. Pillai, S., and Baltimore, D. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 7654–7658
34. Walter, P., and Blobel, G. (1983) Methods Enzymol. 96, 84–93
35. Scheele, G. (1983) Methods Enzymol. 96, 94–111
36. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) Methods Enzymol. 154, 387–391
37. Persa, E., and Lingappa, V. R. (1985) J. Cell Biol. 101, 2292–2301
38. Lammli, U. K. (1970) Nature 227, 680–685
39. Schagger, H., and Von Jagow, G. (1987) Anal. Biochem. 166, 368–379
40. Andrews, D. (1989) Biotechniques 7, 960–967
41. Andrews, D., Lauffer, L., Walter, P., and Lingappa, V. R. (1989) J. Cell Biol. 108, 797–810
42. Tucker, P. W., Marcu, K. B., Newell, N., Richards, J., and Blattner, F. R. (1979) Science 206, 1303–1306
43. Rogers, J., Choi, E., Souza, L., Carter, C., Word, C., Kuehl, M., Eisenberg, D., and Wall, R. (1981) Cell 26, 19–27
44. Yamawaki-Kataoka, Y., Nakai, S., Miyata, T., and Honjo, T. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 2623–2627
45. Cheng, H.-L., Blattner, F. R., Fitzmaurice, L., Mushinski, J. F., and Tucker, P. W. (1982) Nature 296, 410–415
46. Tyler, B., Cowman, A. F., Gerondakis, S. D., Adams, J. M., and Bernard, O. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 2008–2012
47. Thrift, R. N., Andrews, D. W., Walter, P., and Johnson, A. E. (1991) J. Cell Biol. 112, 809–821
48. Eisenberg, D. (1984) Annu. Rev. Biochem. 53, 595–623
49. von Heijne, G. (1987) Protein Sequences & Data Anal. 1, 41–42
50. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
51. Szczesna-Skorupa, E., Browne, N., Mead, D., and Kemper, B. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 738–742
52. Parks, G. D., Hull, J. D., and Lamb, R. A. (1989) J. Cell Biol. 109, 2023–2032