

# The COOH-terminal Ends of Internal Signal and Signal-Anchor Sequences Are Positioned Differently in the ER Translocase

IngMarie Nilsson, Paul Whitley, and Gunnar von Heijne

Karolinska Institute Center for Structural Biochemistry, NOVUM, S-141 57 Huddinge, Sweden

**Abstract.** Signal peptides (SPs) target proteins to the secretory pathway and are cleaved from the nascent chain once the translocase in the ER has been engaged. Signal-anchor (SA) sequences also interact transiently with the ER translocase, but are not cleaved and move laterally out of the translocase to become permanent membrane anchors. One obvious difference between SP and SA sequences is the considerably longer hydrophobic regions (h regions) of the latter. To study the interaction between SP/SA sequences and the ER translocase, we have constructed

signal sequences with poly-Leu h regions ranging in length from 8 to 29 residues and have characterized their locations within the translocase using both a new assay that measures the minimum number of amino acids needed to span the distance between the COOH-terminal end of the h region and the active site of the oligosaccharyl transferase enzyme and an assay where the efficiency of signal peptidase catalyzed cleavage is measured. Our results suggest that SP and SA sequences are positioned differently in the ER translocase.

**I**N eukaryotic cells, most globular and transmembrane proteins destined for compartments along the secretory pathway are initially targeted to a translocation machinery—a “translocase”—in the ER membrane (Gilmore, 1993). The targeting function is provided either by an NH<sub>2</sub>-terminal signal peptide (SP)<sup>1</sup> that is later removed from the polypeptide chain, or by a signal-anchor (SA) sequence that promotes translocation of the part of the chain located on its COOH-terminal side but that is not removed and ends up anchoring the protein in transmembrane orientation (von Heijne, 1988).

When the SP or SA sequence emerges from the ribosome, it first binds to the signal recognition particle (Kurzchalia et al., 1986; Lütcke et al., 1992) and is later transferred to the ER translocase; by chemical cross-linking studies, it has been shown that both SP and SA sequences are located in close proximity to translocase subunits such as Sec61p and TRAM (Görlich and Rapoport, 1993; Hartmann et al., 1994; High et al., 1993a). In mammalian microsomes, other subunits of the translocase include the oligosaccharyl transferase and probably the signal peptidase (Gilmore, 1993; Görlich et al., 1992), while in yeast, a complex consisting of the Sec61p, Sec62p, Sec63p, and Sec66p proteins has been identified (Feldheim et al., 1993).

SP and SA sequences share two common features: a short positively charged NH<sub>2</sub>-terminal segment (n region) and a central hydrophobic stretch (h region). They differ in two important respects, however: the h region is much longer in the SA sequences, and SP sequences end with a short COOH-terminal cleavage region (c region) that contains the signal peptidase cleavage site (von Heijne, 1988). Beyond the demonstration that SP and SA sequences can be cross-linked to the same components of the translocase little is known about the detailed interactions that take place; in particular, it is not known if the difference in h region length has any consequences at the level of SP/SA-translocase interactions, or if it is simply a reflection of the fact that a long h region will provide a more stable membrane anchor once the translocation process has terminated.

A possibility to probe SP/SA-translocase interactions in greater detail was suggested by a previous study where we found that engineered Asn-Ser-Thr glycosylation acceptor sites can only be modified by the oligosaccharyl transferase enzyme when the Asn residue is positioned no closer than a precisely defined “minimum glycosylation distance” of 12–13-residues downstream of an internal, uncleaved SA sequence (Nilsson and von Heijne, 1993). Since the oligosaccharyl transferase is itself an integral part of the ER translocase, we reasoned that its active site could possibly be used as a fixed point of reference against which the position of SP and SA sequences in the translocase can be measured. To test this idea, we have constructed signal sequences with poly-Leu h regions ranging in length from  $n = 8$  to 29 residues. Using *in vitro* translation in the presence of microsomes, we have attempted to define the position of the h region in the

Address all correspondence to G. von Heijne, Karolinska Institute Center for Structural Biochemistry, NOVUM, S-141 57 Huddinge, Sweden. Ph: (46) 8-608 91 31. Fax: (46) 8-774-55-38.

1. *Abbreviations used in this paper:* Lep, leader peptidase; SA, signal-anchor; SP, signal peptide.

translocase by two independent means: (a) by determining the minimum number of amino acids needed to span the distance between the COOH-terminal end of the h region and the first functional acceptor site for N-linked glycosylation as a function of  $n$ ; and (b) by measuring the efficiency of cleavage by signal peptidase as a function of  $n$ . We find a sharp change in the behavior of the signal sequences as  $n$  increases beyond 17–19 residues: signal peptidase cleavage no longer takes place, and the minimum glycosylation distance is reduced by three residues. These observations suggest that SP and SA sequences are positioned differently in the ER translocase.

## Materials and Methods

### Enzymes and Chemicals

Unless otherwise stated, all enzymes were from Promega Biotec (Madison, WI). T7 DNA polymerase was from Pharmacia Fine Chemicals (Piscataway, NJ). Proteinase K was from Merck, Sharpe, and Dohme (Rahway, NJ). [<sup>35</sup>S]Met was from Amersham Corp. (Arlington Heights, IL). Ribonucleotides, deoxyribonucleotides, dideoxyribonucleotides, and the cap analog m<sup>7</sup>G(5')ppp(5')G were from Pharmacia Fine Chemicals. Plasmid pGEM1 and wheat germ lysate was from Promega Biotec. Spermidine, PMSF, BSA, creatine phosphate, and creatine phosphokinase were from Sigma Chemical Co. (St. Louis, MO). The glycosylation acceptor peptide *N*-benzoyl-Asn-Leu-Thr-*N*-methylamide and the nonacceptor peptide *N*-benzoyl-Asn-Leu-(allo)Thr-*N*-methylamide were synthesized according to Erickson and Merrifield (1976). Oligonucleotides were produced using an Applied Biosystems synthesizer 380B followed by NAP-5 (Pharmacia Fine Chemicals) or HPLC purification.

### DNA Techniques

Replacement of the H2 region in Lep by the sequences IK<sub>4</sub>L<sub>*n*</sub>VQ<sub>3</sub>P and IK<sub>4</sub>L<sub>*n*</sub>VPSAQAA was performed by first introducing BclI and NdeI restriction sites in codons 59 and 80 flanking the H2 region and then replacing the BclI-NdeI fragment by the appropriate double-stranded oligonucleotides. Site-specific mutagenesis used to position Asn-Thr-Ser acceptor sites relative to H2 was performed according to the method of Kunkel (Kunkel, 1985), as modified by Geisselsoder et al. (1987). All mutants were confirmed by DNA sequencing of single-stranded M13 DNA using T7 DNA polymerase. For cloning into and expression from the pGEM1 plasmid, the 5' end of the *lep* gene was modified, first, by the introduction of an XbaI site and, second, by changing the context 5' to the initiator ATG codon to a "Kozak consensus" sequence (Kozak, 1989). Thus, the 5' region of the gene was modified to: ...ATAACCTCTAGAGCCACCATGGCGAATATG... (XbaI site and initiator codon underlined). The XbaI-SmaI fragment carrying *lep* was cloned into pGEM1 behind the SP6 promoter.

### In Vitro Transcription and Translation in Reticulocyte Lysate

Synthesis of RNA by SP6 RNA polymerase and translation in reticulocyte lysate in the presence of dog pancreas microsomes was performed as described (Liljeström and Garoff, 1991). Translocation of polypeptides to the luminal side of the microsomes was assayed by resistance to exogenously added proteinase K and by prevention of N-linked glycosylation through competitive inhibition by addition of a glycosylation acceptor tripeptide but not by a nonacceptor tripeptide (Nilsson and von Heijne, 1993).

### In Vitro Translation of Truncated mRNA in Wheat Germ Lysate

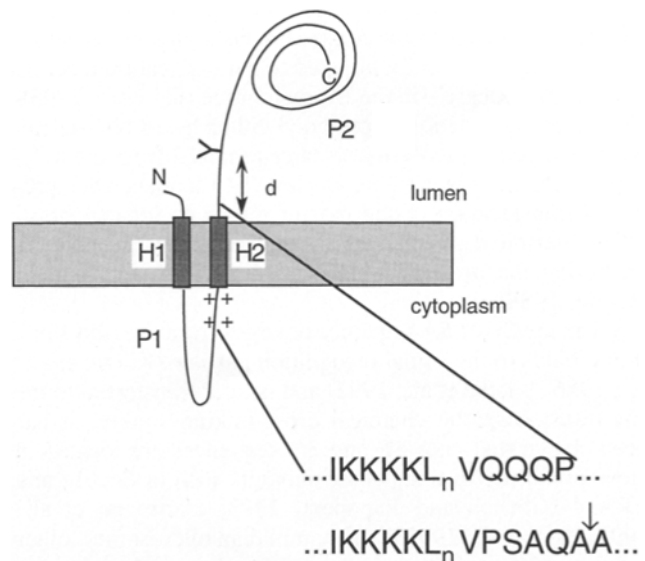
In vitro transcription was performed as described (Liljeström and Garoff, 1991), except that the template DNA was generated by PCR (Miao et al., 1992) and purified from agarose gels using the Magic PCR Prep purification system (Promega Biotec). One PCR primer was designed to hybridize to a region 150 bases upstream of the SP6 promoter in the pGEM1 plasmid and the reverse primers were designed to hybridize within the protein coding region to generate the desired truncation without adding a stop codon.

In vitro translation was performed for 60 min in a 100- $\mu$ l reaction volume using the wheat germ lysate system (Promega Biotec) according to the manufacturers protocol with the addition of 2  $\mu$ l canine SRP (20 nM final concentration) prepared as described (Walter and Blobel, 1983b). The translation mix was supplemented with 4  $\mu$ l (8 A<sub>280</sub> equivalents) EDTA-washed dog pancreas microsomes (Sakaguchi et al., 1992; Walter and Blobel, 1983a) when relevant. At the end of the reaction, samples were acid precipitated by the addition of an equal volume of 20% TCA and then processed for immunoprecipitation with a polyclonal antiserum against Lep.

## Results

### Variation of the Minimum Glycosylation Distance with the Length of the h Region

As a model protein to study the interactions between the ER translocase and SP/SA sequences, we chose the *Escherichia coli* inner membrane protein leader peptidase (Lep) (Fig. 1). The insertion of Lep both into the inner membrane of *E. coli* and into dog pancreas microsomes has been well characterized (Johansson et al., 1993; Nilsson and von Heijne, 1993; von Heijne, 1994): as judged by N-linked glycosylation and protease accessibility, both the NH<sub>2</sub>-terminal tail and the large COOH-terminal P2 domain are efficiently translocated to the luminal side of the microsomal membrane, while the P1 domain remains on the external (cytoplasmic) face of the microsome. Further, we have shown that Asn-Ser-Thr glycosylation acceptor sites (Gavel and von Heijne, 1990) must be located at a minimum distance of 12–13 residues down-

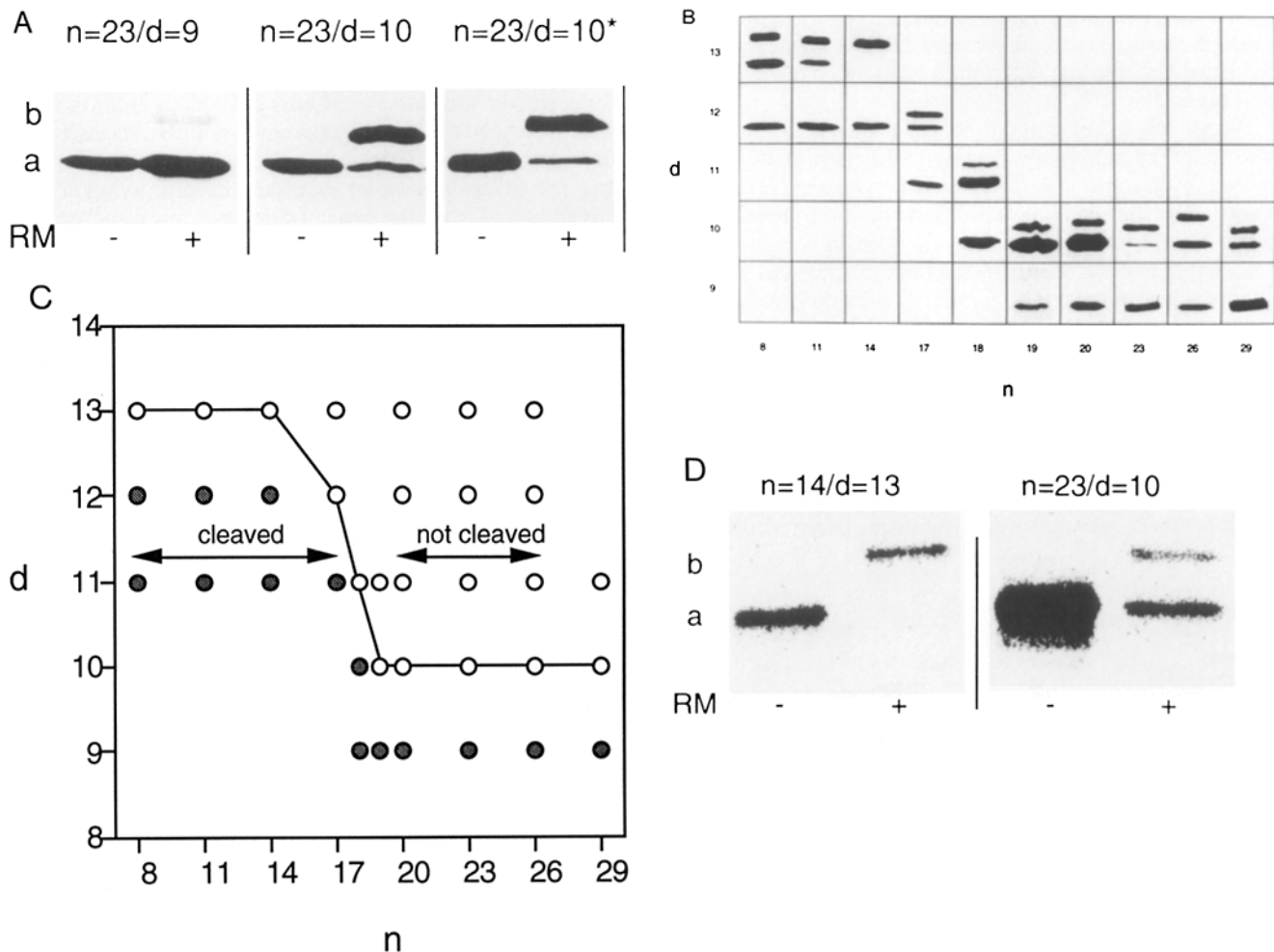


**Figure 1.** Topology of Lep in the microsomal membrane. The H2 region was replaced by the appropriate double-stranded oligonucleotides encoding the amino acid sequence IleLys<sub>4</sub>Leu<sub>*n*</sub>ValGln<sub>3</sub>Pro. For each  $n$ , Asn-Ser-Thr glycosylation acceptor sites were introduced at different positions  $d$  residues downstream of H2 by replacing the corresponding codons in the wild type sequence. A naturally occurring Asn-Glu-Thr glycosylation acceptor site in position 214 in the wild type sequence was changed to the nonacceptor sequence Gln-Glu-Thr in all constructs. In the experiments reported in Fig. 3, the H2 region was replaced by the sequence IleLys<sub>4</sub>Leu<sub>*n*</sub>ValProSerAlaGlnAla<sub>4</sub>Ala, which encodes a predicted COOH-terminal signal peptidase cleavage site (arrow).

stream of the internal SA sequence H2 for glycosylation to be possible (Nilsson and von Heijne, 1993). The object of the present study was to take advantage of this latter observation and use the oligosaccharyl transferase active site as a fixed point of reference and determine the minimum glycosylation distance as a function of the length of the H2 h region.

To avoid possible complications arising from local variations in sequence, H2 was replaced by the sequence Ile-

Lys-Leu-ValGln<sub>3</sub>Pro (h region underlined) (Fig. 1) and  $n$ , the number of leucines in the h region, was varied from 8 to 29 in steps of three. The four lysines on the NH<sub>2</sub>-terminal end of the leucine stretch were included to provide a net positive NH<sub>2</sub>-terminal charge (Sakaguchi et al., 1992; von Heijne, 1984) and a clear demarcation of the start of the h region. For each value of  $n$ , we determined the minimum number of residues  $d_{min}$  needed to span the distance be-



**Figure 2.** Determination of the minimum glycosylation distance  $d_{min}$  as a function of  $n$ , the number of leucines in H2. Synthesis of mRNA from pGEM1 by SP6 RNA polymerase and translation in reticulocyte lysate in the presence of dog pancreas microsomes was performed as described (Liljeström and Garoff, 1991). (A) Results for mutants  $n = 23/d = 9$  (starting from the end of the h region, the relevant sequence reads QQQPYEPFNST),  $n = 23/d = 10$  (sequence QQQPYEPFQNST), and  $n = 23/d = 10^*$  (sequence QQQPYEQPFNST) where  $d$  is the position of the Asn residue in the glycosylation acceptor site relative to the end of the (Leu,Val) segment. Mutant  $n = 23/d = 10^*$  was made from the nonglycosylated mutant  $n = 23/d = 9$  by introducing an extra Gln residue between the h region and the Asn-Ser-Thr acceptor site. Lep mutants were expressed and radiolabeled by [<sup>35</sup>S]Met in the absence (-RM lanes) and presence (+RM lanes) of microsomes. Glycosylation was assayed by including a synthetic glycosylation acceptor tripeptide in the translation mix (Nilsson and von Heijne, 1993) (data not shown). Nonglycosylated (band a) and glycosylated (band b) forms are indicated (Johansson et al., 1993; Nilsson and von Heijne, 1993). (B) Determination of  $d_{min}$ . For each mutant, the result obtained from in vitro translation in the presence of microsomes is shown. The lower band represents nonglycosylated molecules, the upper band glycosylated ones. Results are shown for mutants with a glycosylation distance close to  $d_{min}$ . (C) Summary of glycosylation and signal peptidase cleavage data (cf, Figs. 2 B and 3). Unfilled circles represent constructs that become glycosylated when expressed in the presence of microsomes; filled circles represent nonglycosylated constructs (in addition, constructs with  $d$  values of 15 and 20 were analyzed for  $n = 8, 11, 14, 17, 20,$  and  $23$ ; all were glycosylated). The line shows the variation of  $d_{min}$  with  $n$ . (D) In vitro translation in wheat germ lysate of truncated mRNAs in the absence (-RM) or presence (+RM) of dog pancreas microsomes. Results are shown for the  $n = 14/d = 13$  and the  $n = 23/d = 10$  constructs. The truncations correspond to amino acid position 231 in the wild type leader peptidase sequence which is, respectively, 142- and 145-amino acids downstream of the glycosylation site in the two constructs. The translation products were immunoprecipitated by a polyclonal Lep antiserum, and the gel was visualized on a Fuji phosphorimager.

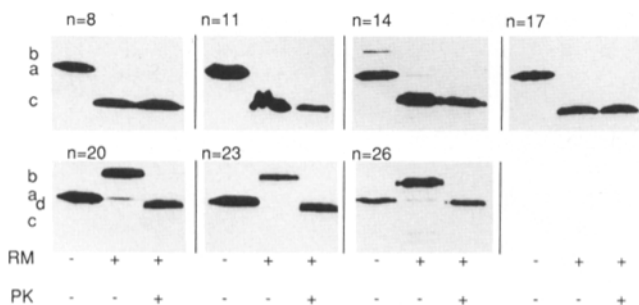
tween the COOH-terminal end of the h region and the Asn residue in the first functional Asn-Ser-Thr acceptor site.

Results for two  $n = 23$  constructs are shown in Fig. 2 A, (left and middle). The transition point from a nonfunctional to a fully functional Asn-Ser-Thr acceptor site can be mapped to within one residue, and  $d_{min} = 10$  residues in this case. As a control, an extra Gln-residue was inserted between the h region and the Asn-Ser-Thr acceptor site in the nonglycosylated ( $n = 23/d = 9$ )-construct to make  $d = 10$ ; as expected, this also led to efficient glycosylation (Fig. 2 A, right), demonstrating that it is the number of residues between the h region and the Asn-Ser-Thr site rather than local sequence features that determines whether or not glycosylation takes place.

Proper formation of the topology depicted in Fig. 1 was also tested by a proteolysis assay (Johansson et al., 1993); indeed, even for  $n = 8$ , the P1 loop between the H1 and H2 transmembrane segments was accessible to proteinase K in intact microsomes whereas the P2 domain was protected, confirming that the (Leu<sub>8</sub>Val) segment functions as a SA sequence (data not shown). Similarly, previous studies have shown that as little as seven contiguous leucine residues can serve as a signal sequence (Sakaguchi et al., 1992).

Results for the full range of constructs studied are shown in Fig. 2 B and are summarized in Fig. 2 C. For  $n = 8$ –14,  $d_{min} = 13$  residues. Over the narrow interval  $n = 17$ –19,  $d_{min}$  drops to 10 residues, and then again remains constant all the way up to  $n = 29$ .

While these results were obtained using full-length



**Figure 3.** Determination of signal peptidase cleavage efficiency as a function of  $n$ . Constructs with a signal peptidase cleavage site after the poly-Leu segment and a glycosylation acceptor site 17-residues downstream of the predicted cleavage site were expressed and radiolabeled by [<sup>35</sup>S]Met in the absence (–RM lanes) and presence (+RM lanes) of microsomes and analyzed by SDS-PAGE. (Band a) Non-glycosylated full-length product. (Band b) Glycosylated full-length product. (Band c) Glycosylated COOH-terminal fragment resulting from signal peptidase cleavage. Translocation into the luminal compartment was also assayed by resistance to proteinase K (Nilsson and von Heijne, 1993) (+ PK lanes); as expected, fragments resulting from cleavage by signal peptidase ( $n \leq 17$ ) were completely resistant to proteolysis while uncleaved constructs ( $n \geq 20$ ) were attacked by proteinase K in the loop between H1 and H2, resulting in a glycosylated, proteinase K-protected fragment (band d). Fortuitously, band d has only a marginally higher mobility than band a. Glycosylation was assayed by including a synthetic glycosylation acceptor tripeptide in the translation mix (Nilsson and von Heijne, 1993); this treatment reduced the apparent size of both the full-length and signal peptidase-cleaved forms, thus demonstrating that they contain functional glycosylation sites (data not shown).

mRNAs translated in a reticulocyte lysate, glycosylation was also tested with truncated mRNAs coding for two constructs,  $n = 14/d = 13$  and  $n = 23/d = 10$ , translated in a wheat germ lysate. In the wheat germ system, the ribosome remains attached to the end of a truncated mRNA, thus preventing disassembly of the translocase complex (Haeupfle et al., 1986; Miao et al., 1992). Under these conditions, a SA sequence can be cross-linked to components of the translocase, demonstrating that it has not yet moved completely out into the surrounding lipid (High et al., 1993a). As seen in Fig. 2 D, the truncated  $n = 14/d = 13$  and  $n = 23/d = 10$  constructs were both glycosylated. As a control, we also expressed a truncated Lep construct where the glycosylation site was placed only 39 codons from the end of the truncated mRNA and thus could not have access to the oligosaccharyl transferase while the nascent chain was still attached to the ribosome; as expected, glycosylation was only observed when the nascent chain was released from the ribosome by treatment with puromycin in this case (data not shown).

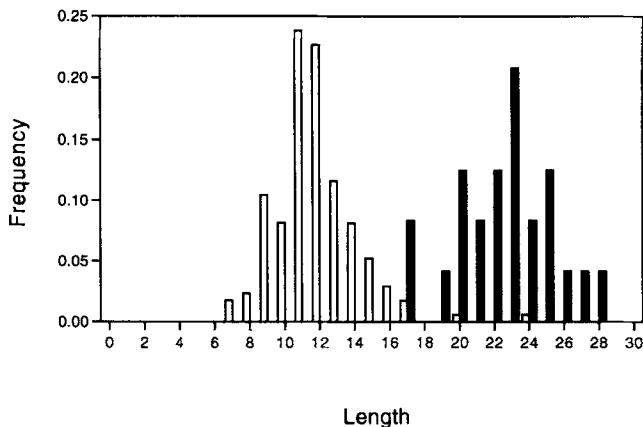
We conclude that the glycosylation reaction is co-translational, and that the COOH-terminal ends of internal signal sequences with short ( $n < 17$ ) and long ( $n \geq 19$ ) h regions are positioned differently relative to the oligosaccharyl transferase active site when the glycosylation reaction takes place.

### The Efficiency of Signal Peptidase Cleavage Is a Function of h Region Length

An independent point of reference for the position of SP sequences in the translocase is provided by the active site of the signal peptidase enzyme (Dalbey and von Heijne, 1992; Gilmore, 1993). We thus inserted a “consensus cleavage cassette” (Nilsson and von Heijne, 1991) immediately downstream of the h region, Fig. 1, thereby converting the uncleaved SA sequence into a potentially cleavable SP sequence. As shown in Fig. 3, cleavage was observed for  $n = 8$ –17, but not for  $n = 20$ –26, again suggesting that the COOH-terminal ends of internal signal sequences with short ( $n \leq 17$ ) and long ( $n \geq 20$ ) h regions are differently positioned in the translocase.

### Discussion

Previously, two methods have been used to study the environment of SP/SA sequences in the ER translocase: light-induced chemical cross-linking (High et al., 1993a,b) and fluorescence quenching of probes attached to specific amino acids in the nascent polypeptide (Crowley et al., 1993). Here, we present the first results obtained with a new technique where the number of residues,  $d_{min}$ , between a defined position in the nascent chain (the COOH-terminal end of the h region) and the oligosaccharyl transferase active site is measured to within one residue. By varying the number of leucine residues in the h region,  $n$ , we show that the distances between the COOH-terminal ends of short ( $n < 17$ ) and long ( $n \geq 19$ ) h regions and the oligosaccharyl transferase active site differ by three residues. Interestingly, a parallel change in the efficiency of signal peptide cleavage is also observed: sequences with short h regions ( $n \leq 17$ ) are efficiently cleaved, while those with long h regions ( $n \geq 20$ ) are not cleaved.



**Figure 4.** h region length distributions for a collection of 172 human signal peptides (□) and 24 human signal-anchor sequences (■) extracted from SwissProt 27.0 (Bairoch and Boeckmann, 1991) using the *Entrez* software (release 8.0; National Center for Biotechnology Information, Bethesda, MD). The h regions were identified visually as stretches of apolar amino acids devoid of charged residues (D, E, K, R) and containing only isolated polar amino acids (S, T, N, Q, P, H). The two signal peptides with h region lengths of 20 and 24 residues both contain multiple Met residues, suggesting that the initiation codons given in the databank may be in error.

Statistical analysis of h regions in eukaryotic SP and SA sequences suggests that a critical length of  $n = 17$ – $18$  residues is not unique to poly-Leu segments. Thus, SP h regions are typically 7–15-residues long, and rarely if ever longer than 17 residues; in contrast, most SA sequences have h regions ranging from 20 to 30 residues in length (von Heijne, 1985, 1986) (Fig. 4). A similar critical length may exist in prokaryotes; indeed, it has been found that a poly-Leu SP is cleaved by the *E. coli* signal peptidase *in vivo* when  $n = 15$  but not when  $n = 20$  (Chou and Kendall, 1990).

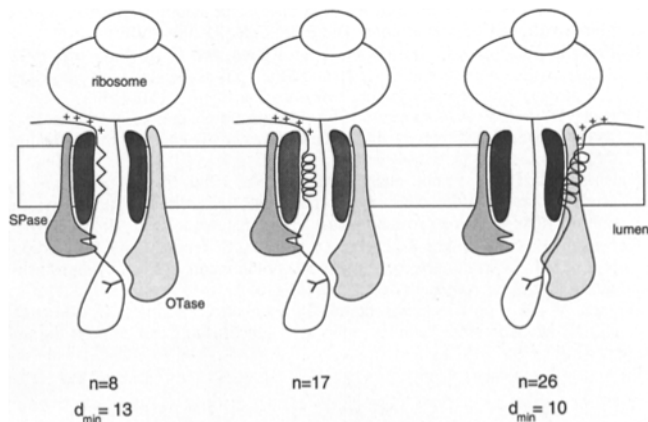
How can we explain the findings that  $d_{min}$  does not vary appreciably with  $n$  in the intervals  $8 \leq n < 17$  and  $19 \leq n \leq 29$ , that  $d_{min}$  is reduced from 13 to 10 residues over the interval  $17 \leq n \leq 19$ , and that cleavage by signal peptidase is only possible for  $n < 20$ ? First, since the sequence around the glycosylation and signal peptidase cleavage sites in our constructs does not vary, we think it unlikely that there are significant conformational differences in the spacer region between the luminal, COOH-terminal end of the h region (the C end) and these sites for different  $n$  values. Considering the amino acid composition of this segment, a flexible, extended conformation seems the most likely. Second, the constancy of  $d_{min}$  over large ranges of  $n$  values means that the position of the C end relative to the oligosaccharyl transferase active site does not vary with  $n$  except in the narrow range  $17 \leq n \leq 19$ , and thus that there are only two possible locations in the translocase for the C end, one for short and one for long h regions. This conclusion is further strengthened by the signal peptidase cleavage results: cleavage is not affected by the length of the h region for  $n \leq 17$ , suggesting that the position of the C end relative to the signal peptidase active site is fixed over this range of  $n$  values, but is completely prevented when  $n \geq 20$ , suggesting that the C end is now in a different location. Since we can only guess at the conformation of the segment between the C end and the

oligosaccharyl transferase active site, all we can say is that the difference in physical distance between the active site and the two different locations of the C end is in the range 5–10 Å (corresponding to three residues in either a helical or a fully extended conformation). A less likely possibility is that the conformation of the spacer segment depends on the location of the C end; if this were the case, the uncertainty in the estimated physical distance will be somewhat greater.

Whether the NH<sub>2</sub>-terminal, cytoplasmic end of the h region (the N end) has a fixed location in the translocase is not known. If in fact this is the case, the conformation of the h region must adapt to changes in  $n$  by being more extended for small  $n$  and becoming increasingly helical as  $n$  increases (Fig. 5, left). This model is rather appealing, as it would equate the “transition region” ( $17 \leq n \leq 19$ ) with the formation of a completely helical h region that would be ideally suited to move into a more peripheral, possibly lipid-exposed location in the translocase. Finally, for larger  $n$ , the lipid-exposed h region would adapt by tilting rather than by exposing hydrophobic residues to the aqueous surroundings (Fig. 5, right). Alternatively, the h region may be helical also for small  $n$ , but this would require that the positively charged lysine residues on its NH<sub>2</sub>-terminal end can be pulled into the translocase.

The present study has been carried out on molecules where the SP/SA sequence is preceded by the transmembrane H1 segment (Fig. 1). Thus, our results pertain to internal SP/SA sequences; whether they also hold for other topogenic sequences such as NH<sub>2</sub>-terminal SP/SA sequences or internal stop-transfer sequences remains to be determined. Since the ER translocase can catalyze the membrane insertion of proteins with multiple transmembrane segments (Lipp et al., 1989; Wessels and Spiess, 1988), it seems likely that each individual transmembrane segment is rapidly moved out of the central channel into a lipid-exposed location as suggested above for SA sequences. If this is true, we should expect to get results very similar to those reported above for both NH<sub>2</sub>-terminal and internal topogenic sequences.

Earlier studies have shown that removal of the n region



**Figure 5.** Possible locations of SP (left and middle) and SA (right) sequences in the ER translocase. The signal peptidase (SPase) and oligosaccharyl transferase (OTase) are shown. Note that the conformation and degree of lipid exposure of the h region is depicted as depending on its length  $n$ . (Adapted from Gilmore, 1993).

from certain SA sequences or point mutations in or near their h regions can uncover cryptic signal peptidase cleavage sites (Hegner et al., 1992; Lipp and Dobberstein, 1986; Roy et al., 1993; Schmid and Spiess, 1988) suggesting that residues in regions flanking the h region can influence the position of SA sequences in the translocase; such effects can now be directly studied with the glycosylation mapping technique described here. The technique may also be used to characterize the interactions between the translocase and h regions of varying degrees of hydrophobicity and to measure other distances in the translocase; such measurements are in progress.

Dog pancreas microsomes were a gift from Dr. Masao Sakaguchi (Kyushu University, Fukuoka, Japan). Glycosylation acceptor and nonacceptor tripeptides were kindly provided by Dr. Henrik Garoff (Karolinska Institute Center for Biotechnology, Huddinge, Sweden). Canine SRP was a gift from Dr. Tom Rapoport (Max-Delbrück Center for Molecular Medicine, Berlin, Germany). This work was supported by grants from the Swedish Natural Sciences Research Council (NFR), the Swedish Technical Research Council (TFR), and the Swedish National Board for Industrial and Technical Development (NUTEK) to G. von Heijne.

Received for publication 25 February 1994 and in revised form 10 June 1994.

## References

- Bairoch, A., and B. Boeckmann. 1991. The SwissProt protein sequence data bank. *Nucleic Acids Res.* 19:2247-2249.
- Chou, M. M., and D. A. Kendall. 1990. Polymeric sequences reveal a functional interrelationship between hydrophobicity and length of signal peptides. *J. Biol. Chem.* 265:2873-2880.
- Crowley, K. S., G. D. Reinhart, and A. E. Johnson. 1993. The signal sequence moves through a ribosomal tunnel into a noncytoplasmic aqueous environment at the ER membrane early in translocation. *Cell.* 73:1101-1115.
- Dalbey, R. E., and G. von Heijne. 1992. Signal peptidases in prokaryotes and eukaryotes: a new protease family. *Trends Biochem. Sci.* 17:474-478.
- Erickson, B. W., and R. B. Merrifield. 1976. Solid phase peptide synthesis. In *The Proteins*. R. L. Hill and H. Neurath, editors. Academic Press, London. 255-257.
- Feldheim, D., K. Yoshimura, A. Admon, and R. Schekman. 1993. Structural and functional characterization of Sec66p: a new subunit of the polypeptide translocation apparatus in the yeast endoplasmic reticulum. *Mol. Biol. Cell.* 4:931-939.
- Gavel, Y., and G. von Heijne. 1990. Sequence differences between glycosylated and non-glycosylated Asn-X-Thr Ser acceptor sites: implications for protein engineering. *Protein Eng.* 3:433-442.
- Geisselsoder, J., F. Witney, and P. Yuckenberg. 1987. Efficient site-directed in vitro mutagenesis. *Biotechniques.* 5:786-791.
- Gilmore, R. 1993. Protein translocation across the endoplasmic reticulum—a tunnel with toll booths at entry and exit. *Cell.* 75:589-592.
- Görlich, D., S. Prehn, E. Hartmann, K. U. Kalies, and T. A. Rapoport. 1992. A mammalian homolog of SEC61p and SECYp is associated with ribosomes and nascent polypeptides during translocation. *Cell.* 71:489-503.
- Görlich, D., and T. A. Rapoport. 1993. Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell.* 75:615-630.
- Haeuptle, M. T., R. Frank, and B. Dobberstein. 1986. Translation arrest by oligodeoxynucleotides complementary to mRNA coding sequences yields polypeptides of predetermined length. *Nucleic Acids Res.* 14:1427-1448.
- Hartmann, E., T. Sommer, S. Prehn, D. Görlich, S. Jentsch, and T. A. Rapoport. 1994. Evolutionary conservation of components of the protein translocation complex. *Nature (Lond.).* 367:654-657.
- Hegner, M., A. von Kieckebuschguck, R. Falchetto, P. James, G. Semenza, and N. Mantei. 1992. Single amino acid substitutions can convert the un-cleaved signal-anchor of sucrose-isomaltase to a cleaved signal sequence. *J. Biol. Chem.* 267:16928-16933.
- High, S., S. S. L. Andersen, D. Görlich, E. Hartmann, S. Prehn, T. A. Rapoport, and B. Dobberstein. 1993a. Sec61p is adjacent to nascent type-I and type-II signal-anchor proteins during their membrane insertion. *J. Cell Biol.* 121:743-750.
- High, S., B. Martoglio, D. Görlich, S. S. L. Andersen, A. J. Ashford, A. Giner, E. Hartmann, S. Prehn, T. A. Rapoport, B. Dobberstein, and J. Brunner. 1993b. Site-specific photocross-linking reveals that Sec61p and TRAM contact different regions of a membrane-inserted signal sequence. *J. Biol. Chem.* 268:26745-26751.
- Johansson, M., I. Nilsson, and G. von Heijne. 1993. Positively charged amino acids placed next to a signal sequence block protein translocation more efficiently in *Escherichia coli* than in mammalian microsomes. *Mol. Gen. Genet.* 239:251-256.
- Kozak, M. 1989. Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol. Cell Biol.* 9:5073-5080.
- Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA.* 82:488-492.
- Kurzchalia, T. V., M. Wiedmann, A. S. Girshovich, E. S. Bochkareva, H. Bielka, and T. A. Rapoport. 1986. The signal sequence of nascent preprolactin interacts with the 54K polypeptide of the signal recognition particle. *Nature (Lond.).* 320:634-636.
- Liljeström, P., and H. Garoff. 1991. Internally located cleavable signal sequences direct the formation of semliki forest virus membrane proteins from a polyprotein precursor. *J. Virol.* 65:147-154.
- Lipp, J., and B. Dobberstein. 1986. The membrane-spanning segment of invariant chain (I gamma) contains a potentially cleavable signal sequence. *Cell.* 46:1103-1112.
- Lipp, J., N. Flint, M.-T. Haeuptle, and B. Dobberstein. 1989. Structural requirements for membrane assembly of proteins spanning the membrane several times. *J. Cell Biol.* 109:2013-2022.
- Lütcke, H., S. High, K. Römisch, A. J. Ashford, and B. Dobberstein. 1992. The methionine-rich domain of the 54 kDa subunit of signal recognition particle is sufficient for the interaction with signal sequences. *EMBO (Eur. Mol. Biol. Organ.) J.* 11:1543-1551.
- Miao, G.-H., Z. Hong, and D. P. S. Verma. 1992. Topology and phosphorylation of soybean Nodulin-26, an intrinsic protein of the peribacteroid membrane. *J. Cell Biol.* 118:481-490.
- Nilsson, I. M., and G. von Heijne. 1991. A *de novo* designed signal peptide cleavage cassette functions *in vivo*. *J. Biol. Chem.* 266:3408-3410.
- Nilsson, I., and G. von Heijne. 1993. Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. Biol. Chem.* 268:5798-5801.
- Roy, P., C. Chatellard, G. Lemay, P. Crine, and G. Boileau. 1993. Transformation of the signal peptide/membrane anchor domain of a type-II transmembrane protein into a cleavable signal peptide. *J. Biol. Chem.* 268:2699-2704.
- Sakaguchi, M., R. Tomiyoshi, T. Kuroiwa, K. Mihara, and T. Omura. 1992. Functions of signal and signal-anchor sequences are determined by the balance between the hydrophobic segment and the N-terminal charge. *Proc. Natl. Acad. Sci. USA.* 89:16-19.
- Schmid, S. R., and M. Spiess. 1988. Deletion of the amino-terminal domain of asialoglycoprotein receptor H1 allows cleavage of the internal signal sequence. *J. Biol. Chem.* 263:16886-16891.
- von Heijne, G. 1984. Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in prokaryotic and eukaryotic cells. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:2315-2318.
- von Heijne, G. 1985. Signal sequences. The limits of variation. *J. Mol. Biol.* 184:99-105.
- von Heijne, G. 1986. Towards a comparative anatomy of N-terminal topogenic protein sequences. *J. Mol. Biol.* 189:239-242.
- von Heijne, G. 1988. Transcending the impenetrable: how proteins come to terms with membranes. *Biochim. Biophys. Acta.* 947:307-333.
- von Heijne, G. 1994. Membrane proteins: from sequence to structure. *Ann. Rev. Biophys. Biomol. Struct.* 23:167-192.
- Walter, P., and G. Blobel. 1983a. Preparation of microsomal membranes for cotranslational protein translocation. *Methods Enzymol.* 96:84-93.
- Walter, P., and G. Blobel. 1983b. Signal recognition particle: a ribonucleoprotein required for cotranslational translocation of proteins, isolation and properties. *Methods Enzymol.* 96:682-691.
- Wessels, H. P., and M. Spiess. 1988. Insertion of a multispanning membrane protein occurs sequentially and requires only one signal sequence. *Cell.* 55:61-70.