THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 9, Issue of March 3, pp. 6207–6213, 2000

Printed in U.S.A.

Distant Downstream Sequence Determinants Can Control N-tail Translocation during Protein Insertion into the Endoplasmic Reticulum Membrane*

(Received for publication, August 27, 1999, and in revised form, December 9, 1999)

IngMarie Nilsson‡, Susanne Witt¶, Hans Kiefer§, Ismael Mingarro¶, and Gunnar von Heijne‖

From the Department of Biochemistry, Stockholm University, S-10691 Stockholm, Sweden

We have studied the membrane insertion of ProW, an *Escherichia coli* inner membrane protein with seven transmembrane segments and a large periplasmic N-terminal tail, into endoplasmic reticulum (ER)-derived dog pancreas microsomes. Strikingly, significant levels of N-tail translocation is seen only when a minimum of four of the transmembrane segments are present; for constructs with fewer transmembrane segments, the N-tail remains mostly nontranslocated and the majority of the molecules adopt an "inverted" topology where normally nontranslocated parts are translocated and vice versa. N-tail translocation can also be promoted by shortening of the N-tail and by the addition of positively charged residues immediately downstream of the first trasnmembrane segment. We conclude that as many as four consecutive transmembrane segments may be collectively involved in determining membrane protein topology in the ER and that the effects of downstream sequence determinants may vary depending on the size and charge of the N-tail. We also provide evidence to suggest that the ProW N-tail is translocated across the ER membrane in a C-to-N-terminal direction.

The structure and function of integral membrane proteins are in part determined by their topology, *i.e.* the orientation (N_{cyt} or N_{exo}) of each transmembrane segment (TM)[1] relative to the membrane. Two general mechanisms for the membrane assembly of a polytopic membrane protein have been proposed. In the first, the overall topology is assumed to be determined by the most N-terminal TM, with downstream TMs serving alternately as stop transfer and signal anchor sequences (1). The second model suggests that topogenic information is spread throughout the protein and that downstream TMs may affect the orientation of upstream ones (2–4).

One aspect of membrane protein assembly that is not fully understood is the translocation of N-terminal tails (N-tails) across the membrane (5). N-tail translocation is presumably initiated by a hydrophobic "reverse signal-anchor" sequence that also becomes the first TM segment (6–10). In eukaryotic cells, N-tail translocation has been shown to proceed by the normal signal recognition particle-Sec61 pathway (11) and to require an unfolded N-terminal domain (12). In contrast, N-tail translocation in *Escherichia coli* appears to be possible both by a SecA-dependent (10) and by a Sec-independent mechanism (6–8, 13), depending on the protein. In both prokaryotic and eukaryotic cells, it seems that the presence of too many positively charged amino acids can prevent N-tail translocation, whereas negatively charged residues may facilitate translocation (6, 7, 14–17).

The 100-residue-long periplasmic N-terminal domain of the *E. coli* inner membrane protein ProW (6, 18) is one of the longest known N-tails. Earlier studies in *E. coli* have shown that the ProW N-tail can be efficiently translocated across the inner membrane (6, 7). We have now expressed ProW *in vitro* in the presence of ER-derived dog pancreas microsomes and have studied N-tail translocation in the full-length protein as well as in fusion constructs lacking one or more TM segments and with N-tails of different lengths. Strikingly, and in contrast to previous findings in *E. coli*, we find that a minimum of four TM segments must be present to reach appreciable levels of N-tail translocation; constructs with up to three TM segments mainly adopt an "inverted" topology with the N-tail in the cytoplasm. Shortening of the N-tail increases translocation efficiency and reduces the number of TMs that need to be present for N-tail translocation to occur.

The full-length ProW N-tail is efficiently translocated across the ER membrane when fused to the N terminus of another *E. coli* inner membrane protein, leader peptidase (Lep). Studies of the kinetics of glycosylation of Asn-Xaa-Thr acceptor sites located either early or late in the N-tail in this construct strongly suggest that N-tail translocation proceeds in a C-to-N-terminal direction.

These observations show that the whole TM1-TM4 region of ProW can influence N-tail translocation across the ER membrane. Thus, distant downstream sequence determinants can affect the insertion of the most N-terminal parts of a polytopic membrane protein into the ER membrane.

MATERIALS AND METHODS

Enzymes and Chemicals—Unless otherwise stated, all enzymes as well as plasmid pGEM1, RiboMAX SP6 RNA polymerase system, and rabbit reticulocyte lysate were from Promega (Madison, WI) or New England Biolabs (Boston, MA). T7 DNA polymerase, *Taq* polymerase, [^{35}S]Met and ^{14}C-methylated marker proteins, ribonucleotides, deoxyribonucleotides, dideoxyribonucleotides, and the cap analog m7G(5′)ppp(5′)G were from Amersham Pharmacia Biotech. Aurintricarboxylic acid was from Sigma. Qiagen PCR purification kit and Qiagen RNeasy RNA clean up were from Qiagen (Hilden, Germany). Oligonucleotides were from Kebo Lab and Cybergene (Stockholm, Sweden).

DNA Techniques—Site-specific mutagenesis was performed accord-

* This work was supported by grants from the Swedish Natural and Technical Sciences Research Councils, the Swedish Cancer Foundation, and the Göran Gustafsson Foundation and by European Commission Grant BIO4-CT96-0129 (to G. v. H.). 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.

‡ These authors contributed equally to this work.

§ Present address: Inst. of Microbiology, University of Hohenheim, Garbenstrasse 30, D-70599 Stuttgart, Germany.

¶ Present address: Dept. of Biochemistry, University of Valencia, Dr. Moliner 50, E-46100 Burjossot, Valencia, Spain.

‖ To whom correspondence should be addressed. Tel.: 46-8-16-25-90; Fax: 46-8-15-36-79; E-mail: gunnar@biokemi.su.se.

[1] The abbreviations used are: TM, transmembrane segment; Lep, leader peptidase; PCR, polymerase chain reaction.

ing to the method of Kunkel (19, 20) or by PCR. All mutants were confirmed by sequencing of plasmid DNA. PCR was used to amplify fragments from pING1 (21) and pGEM1 plasmids containing the desired DNA constructs. The amplified DNA products were purified using Qiagen PCR purification kit as described in the manufacturer's protocol. All cloning steps were done according to standard procedures.

Construction of Full-length and Truncated ProW Glycosylation Mutants—The gene coding for *E. coli* ProW was amplified by PCR from genomic DNA using 5′- and 3′-specific oligonucleotides containing N- or C-terminal glycosylation sites (see below) and suitable restriction sites for cloning into pGEM1. The full-length ProW construct containing an N-terminal glycosylation site was engineered using a 5′-specific oligonucleotide, which introduced a mutation encoding the *N*-glycosylation site Asn^5-Asn^6-Thr^7 near the N terminus of ProW, a Kozak consensus sequence (22) for enhanced translation, and a *Xba*I site for cloning (modified nucleotides underlined): . . . ACCTCTAGAGCCACCATG-GCTGATCAAAATAATACGTGGGATACCACGCCAGCG. . . . The reverse oligonucleotide encoded the 3′ end of ProW, a stop codon, and a *Sma*I site for cloning. Truncated ProW molecules were made in the same way, but with the reverse 3′ oligonucleotide (including stop codons) hybridizing to the relevant internal portions of the *proW* gene.

The C-terminal C* reporter domain was introduced by a two-step PCR amplification. In the first step a common 5′-specific primer and individual 3′-primers specific for the full-length gene and the various truncations, were used. The common 5′-specific primer was identical to the one described above except that it lacked the glycosylation acceptor site. The individual 3′-oligonucleotides introduced a C-terminal Asn-Leu-Thr glycosylation site located at least 16 amino acid residues downstream of the nearest transmembrane domain. To increase the glycosylation efficiency of the C* reporter, a second PCR amplification was done using the same common 5′-specific primer described above and a common 3′-specific primer introducing a C-terminal extension of 18 amino acid residues downstream of the Asn-Leu-Thr site followed by a *Sma*I restriction site for cloning into pGEM1. The encoded sequence of the spacer behind the glycosylation site was SGKENGIRLSER-KETLGD. The PCR products were cloned into *Xba*I-*Sma*I-restricted pGEM1 plasmids.

Construction of ProW/P2 and ProW/P2′ Fusions—The construction of ProW/P2 fusions in pING1 containing up to three transmembrane segments was described previously (23). Cloning into pGEM1 was done using ProW/P2 *Xba*I-*Sma*I fragments that were amplified from the corresponding pING1 constructs using a universal 5′-specific primer containing an *Xba*I site and a Kozak consensus sequence (see above) and individual 3′-specific primers containing a *Sma*I site. For the construction of ProW/P2 fusions containing four to seven transmembrane segments, the appropriate fragments amplified by PCR from the genomic copy of ProW were cloned into a *Xba*I-*Kpn*I-restricted pGEM1 ProW(TM1–3)/P2 plasmid. The ProW/P2′ fusions were constructed in a similar way, except that the 3′ primer introduced a stop codon in position 216 in Lep (this removes the naturally occurring glycosylation site at Asn^{214}), together with a *Sac*I restriction site. Site-specific mutagenesis was used to introduce three Arg residues between positions 119 and 120 in the ProW TM1–2 loop. The three Arg mutants were cloned into pGEM1 as an *Xba*I-*Sma*I fragment.

Shortening of the N-tail in ProW/P2 and ProW/P2′ Fusions—mRNAs encoding constructs with shortened ProW N-tails were prepared using PCR to amplify fragments from the relevant pGEM1 plasmids.[2] The 5′ primers all had the common sequence 5′-GATTTAGGTGACACTATAGAGGAAACAGGACCATGGCCAATTCCA-CC. . . . .-3′ and contained the SP6 transcriptional promoter, a ribosome-binding site, an initiator codon, and a Asn-Ser-Thr glycosylation site (underlined). The unique sequence at the 3′ end was designed to hybridize at the position of the first residue after the intended deletion in the N-tail. Constructs with ProW residues 1–30, 1–50, and 1–70 deleted were made (note that the PCR primer adds 5 residues to the N terminus of the deletion constructs). The 3′ primers were chosen to have a stop codon in position 216 or 324 in Lep as detailed above.

Construction of ProW N-tail-Lep Fusions—Nt/Lep fusions were constructed by PCR amplification of relevant fragments of the *proW* and *lep* genes to create in-frame fusions with Lep residues 1–323 or 1–215. The fusion joint in the Nt/Lep fusion proteins was -QQ^{99} TRM^1A- (numbers refer to the wild-type ProW and Lep sequences). The ProW N-tail was amplified using a 5′ primer situated 210 bases upstream of the translation start and a 3′ primer at position 99 in ProW containing a *Spe*I site. The 5′ primer for the amplification of Lep contained a *Xba*I site and the 3′ primer had a stop codon in positions 216 or 324. The cleaved PCR fragments were purified on agarose gel and were ligated directly in the gel. Templates for *in vitro* transcription of mRNA were prepared using PCR to amplify the ligated fragments using the same SP6 promoter-containing 5′ primer and the same 3′ primers with stop codons at either position 216 or 324 in Lep as above. In construct *Nt(Asn^{80})/Lep-P2′, the glycosylation acceptor site in the ProW N-tail had the sequence Asn^{80}-Ser-Thr.

In Vitro Transcription—Templates for *in vitro* transcription were prepared as described in Ref. 24 or by PCR amplification with the pGEM1 constructs as template. The 5′ primer was the same in all cases and had the sequence 5′-TTCGTCCAACCAAACCGACTC-3′ (except when the SP6 promoter-containing primers were used; see above). This primer is situated 210 bases upstream of the translational start, and the amplified fragments thus contained the SP6 transcriptional promoter from pGEM1. The 3′ primers contained stop codons in the appropriate positions. Amplified PCR fragments were transcribed from the SP6 promoter using the Large Scale RNA Synthesis kit with the RiboMAX SP6 RNA polymerase system. Transcriptions were carried out at 30 °C for 12 h. The mRNAs were purified using Qiagen RNeasy clean up kit and verified on a 1% agarose gel.

In Vitro Translation—Translation in reticulocyte lysate in the presence of dog pancreas microsomes was performed as described in Ref. 24. Sodium carbonate extraction of microsomes was carried out as described in Ref. 25. Proteins were analyzed by SDS-polyacrylamide gel electrophoresis, and gels were quantitated on a Fuji BAS1000 phosphoimager using the MacBAS 2.31 software. The extent of glycosylation of a given mutant was calculated as the quotient between the intensity of the glycosylated band divided by the summed intensities of the glycosylated and nonglycosylated bands. In general, the glycosylation efficiency varied by no more than ± 5% between repeat experiments. Kinetics of glycosylation was measured as described in Ref. 26, and the results were quantitated by phosphoimager analysis.

RESULTS

To study N-tail translocation in the ProW protein in a microsome-supplemented *in vitro* translation system, we engineered series of ProW fusion proteins and truncated ProW variants containing unique N- or C-terminal N-glycosylation acceptor Asn-Xaa-Thr sites. In most constructs, the N-terminal glycosylation site was placed at residues 5–7 in the ProW N-tail by mutating these residues to Asn-Asn-Thr. As C-terminal reporters we used both the P2-domain (residues 81–323) from the *E. coli* inner membrane protein Lep and a short C-terminal tag (C*) including an Asn-Leu-Thr glycosylation acceptor site followed by an 18-residue-long tail (see "Materials and Methods"). The P2 reporter has a naturally occurring glycosylation site (Asn^{214}-Glu-Thr) and is well suited for topological mappings based on *N*-linked glycosylation (27, 28). A truncated version of the P2 domain, P2′ (residues 81–215), lacking the glycosylation site was also used. All constructs were expressed *in vitro* in the absence or presence of dog pancreas rough microsomes. Because *N*-linked glycosylation occurs only in the lumen of the microsomes, the localization of the introduced acceptor sites can be determined by assaying their glycosylation status. Addition of a single *N*-linked oligosaccharide to the nascent chain leads to an increase in molecular mass of about 2 kDa that is easily detectable by SDS-polyacrylamide gel electrophoresis.

Topology of Full-length ProW in Microsomes—The topology of ProW in the inner membrane of *E. coli* has previously been determined by PhoA/LacZ fusion analysis and protease mapping (6, 18). As shown in Fig. 1, the protein has seven TM segments, and the long N-tail is located in the periplasm. The topology of full-length ProW in microsomes was probed by (i) introducing a glycosylation acceptor site in positions 5–7 in the N-tail, (ii) constructing a full-length ProW/P2* fusion, and (iii) adding a short C-terminal extension (C*) containing a glycosylation acceptor site to the full-length protein. In what follows, an asterisk always indicates the presence of an Asn-Xaa-Thr glycosylation acceptor site in the relevant domain.

As shown in Fig. 2, ~30% of the molecules with the N-

[2] W. Mothes, personal communication.

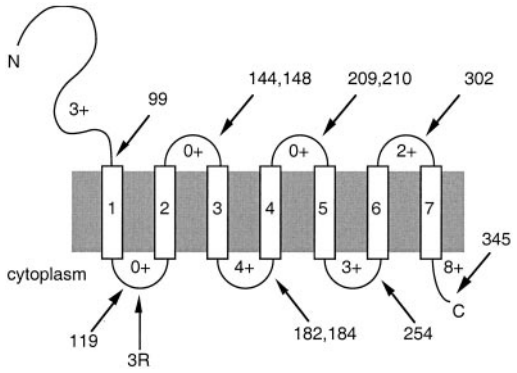

FIG. 1. **Topology of ProW in *E. coli* (6, 18).** The positions of all P2, P2′, and C* fusions and C-terminal truncations are indicated (when two numbers are given, the first is for the P2 and P2′ fusions). The position of the three-Arg insertion (*3R*) in the TM1–2 loop and the number of positively charged residues (Arg+Lys) in the different loops are also shown.

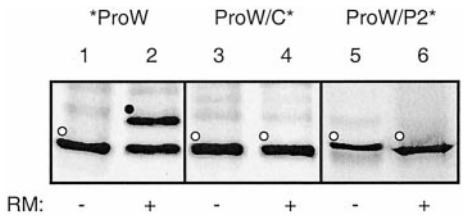

FIG. 2. **Localization of the N and C termini of full-length ProW relative to the microsomal membrane.** *ProW (*lanes 1* and *2*), ProW/C* (*lanes 3* and *4*), and ProW/P2* (*lanes 5* and *6*) (*asterisk* indicates the location of a unique glycosylation site in the relevant domain) were expressed *in vitro* in the absence (−) and presence (+) of rough dog pancreas microsomes (*RM*). Nonglycosylated and glycosylated forms are indicated by *white* and *black dots*, respectively.

terminal acceptor site were glycosylated when expressed *in vitro* in the presence of microsomes, whereas essentially no glycosylation was seen for the two C-terminal reporters. Because the maximal glycosylation efficiency routinely obtained in our *in vitro* system is ~80% (see, *e.g.* construct TM1/P2* in Fig. 3*A* and construct *TM1–3(3R)/P2′ in Fig. 5*A*), we conclude that the N-tail is translocated into the lumen of the microsomes in ~40% of the molecules and that the ProW C terminus is on the cytoplasmic side of the membrane.

Topology of Truncated ProW Constructs—To map the topology of ProW in more detail, we constructed fusions where the reporter domains (the P2* and P2′ domains and the C* tag) were fused at various locations in ProW (Fig. 1).

The results for fusions to the P2* domain are summarized in Fig. 3*A* (*black bars*). Unexpectedly, the P2* fusions after TM1 and TM3 were efficiently glycosylated, whereas the fusion after TM2 was not glycosylated. On the other hand, and in agreement with the topology determined in *E. coli*, the fusions after TM4 and TM6 were more efficiently glycosylated than those after TM5 and TM7. These observations suggest that the TM1–3 part of ProW, when expressed without the downstream TM segments, inserts with an inverted topology compared with that adopted by the full-length molecules. In contrast, the majority of the molecules in the longer ProW/P2* fusions have the same C-terminal orientation as in *E. coli*.

To check whether the length of the C-terminal fusion domain could affect the topology, fusions to the short C* tag were also analyzed (Fig. 3*A*, *white bars*). In general, glycosylation levels were lower in this case; this was expected, because glycosylation acceptor sites located close to the C terminus of a protein

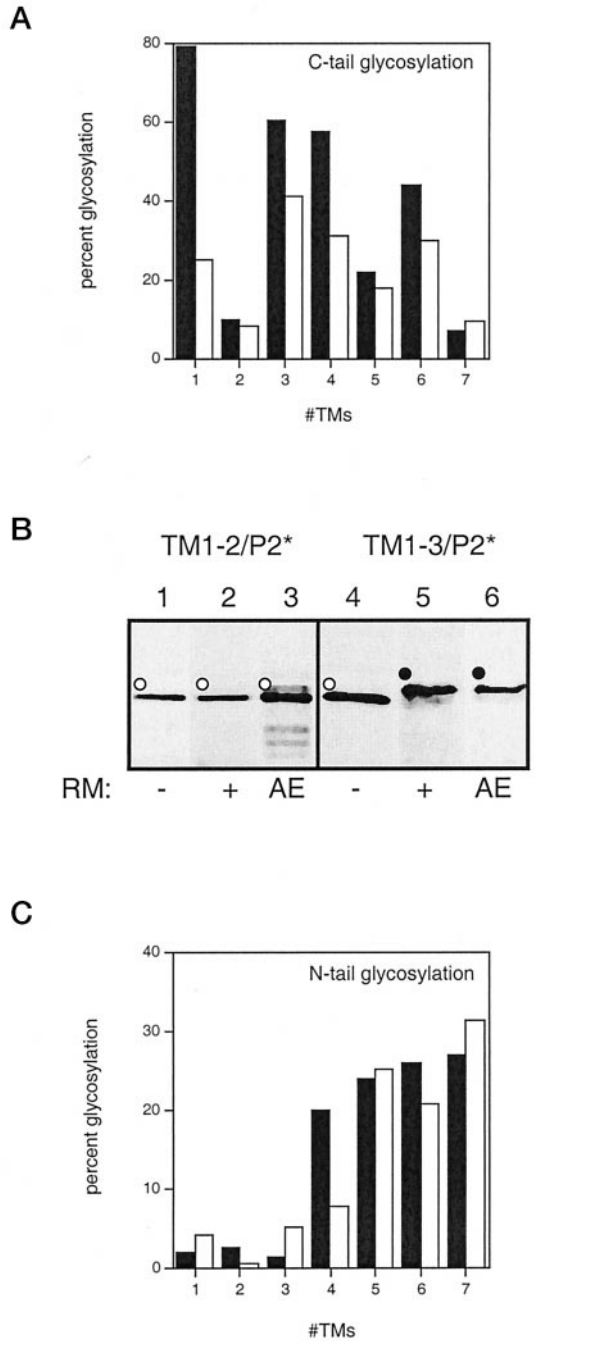

FIG. 3. **Topology mapping of truncated ProW constructs.** *A*, degree of glycosylation of the P2* and C* C-terminal reporter domains in ProW/P2* fusions (*black bars*) and ProW/C* fusions (*white bars*). *B*, alkaline extraction of ProW/P2* fusions. Constructs were expressed *in vitro* in the absence (−) and presence (+) of rough microsomes (*RM*), and the microsomes were subjected to alkaline extraction (*AE*) before loading onto the gel. *Lanes 1–3*, TM1–2/P2*; *lanes 4–6*, TM1–3/P2*. Nonglycosylated and glycosylated forms are indicated by *white* and *black dots*, respectively. *C*, degree of glycosylation of the N-tail in *ProW/P2′ fusions (*black bars*) and C-terminally truncated *ProW constructs (*white bars*).

are less efficiently modified than internal sites (29).[3] The results for the C* fusions paralleled those of the P2* fusions, except for construct TM1/C*, which was poorly glycosylated. Possibly, because co-translational targeting may not be possible in this construct where the TM1 segment is still mostly inside the ribosome at the time of chain termination, the pro-

[3] I. Nilsson and G. von Heijne, unpublished data.

tein is only inefficiently targeted to the microsomes. In general, however, the qualitative picture is the same for the ProW/P2* and ProW/C* fusions, suggesting that the length of the C-terminal reporter domain has little influence on the topology.

To rule out the possibility that lack of glycosylation was a consequence of lack of membrane insertion of the protein, we checked the membrane insertion of the TM1–2/P2* and TM1–3/P2* constructs by carbonate extraction (30) (Fig. 3*B*). Both constructs were quantitatively retained in the alkali-extracted membrane pellet, demonstrating proper assembly into the microsomal membrane.

Because the C-terminal fusions suggested that the topology may be different for proteins with up to three TM domains and those with four or more, we also assayed N-tail translocation in these constructs. To this end, we expressed *ProW/P2′ fusions and truncated *ProW molecules with a unique glycosylation acceptor site in the N-tail (Fig. 3*C*). Little glycosylation of the N-tail was seen for the *TM1/P2′, *TM1–2/P2′, and *TM1–3/P2′ constructs, whereas the *TM1–4/P2′, *TM1–5/P2′, *TM1–6/P2′, and *TM1–7/P2′ constructs were glycosylated to between 20 and 30% (*black bars*). Similar results were seen with the truncated ProW molecules (*white bars*), except that the *TM1–4 construct was very inefficiently glycosylated. Thus, N-tail translocation increases significantly when at least four (P2′ fusions) or five (truncated ProW molecules) TM segments are present. The length of the C-terminal tail thus makes little difference, except for the *TM1–4 constructs.

We conclude that only a very minor fraction of the ProW/P2 constructs with up to three TM segments is oriented with the N-tail in the lumen, whereas a significant fraction (30–40%) of those with four or more TM segments has a lumenal N-tail. The efficiency of N-tail translocation in the TM1–4 constructs apparently depends on the length of the C-terminal tail, with a longer C-tail favoring a lumenal orientation of the N-tail.

Addition of Positively Charged Residues in the TM1-TM2 Loop Affects the Topology of the TM1–2/P2 and TM1–3/P2 Constructs—The TM1–2 loop lacks positively charged residues (Fig. 1), which may contribute to the unexpected topology of the shorter ProW constructs. To test this notion, three positively charged arginine residues were inserted between positions 119 and 120 in the TM1-TM2 loop, and the topology of the resulting TM1–2(3R) and TM1–3(3R) P2* and P2′ fusions was probed by the glycosylation status of unique N-terminal and C-terminal acceptor sites (Fig. 4*A*). The N-tail was still poorly glycosylated in both the *TM1–2(3R)/P2′ and *TM1–3(3R)/P2′ constructs. In contrast, the P2-domain was more efficiently glycosylated in the TM1–2(3R)/P2* construct than in the parent TM1–2/P2* construct (34% *versus* 10%; *cf.* Fig. 3*A*), while the TM1–3(3R)/P2* construct was much less efficiently glycosylated than its parent (7% *versus* 60%). To rule out the possibility that the position of the glycosylation site in the N-tail affected these results, we introduced a second acceptor site in position 80, *i.e.* 20 residues upstream of TM1. The two resulting constructs, **TM1–3/P2′ and **TM1–3(3R)/P2′, were as poorly glycosylated as the parent constructs (data not shown). We also introduced six consecutive arginines in the TM1-TM2 loop; again, no N-tail translocation was seen for the resulting *TM1–3(6R)/P2′ construct (data not shown). Thus, although the introduction of positively charged residues in the TM1-TM2 loop leads to its retention in the cytoplasm (and translocation of the TM2-TM3 loop to the lumen), it does not promote translocation of the N-tail (Fig. 4*B*).

Shortening of the N-tail Increases Translocation Efficiency—To study the effect of the length of the N-tail on its translocation, we deleted 24, 45, and 65 N-terminal amino acid residues in the *TM1/P2′, *TM1–2/P2′, *TM1–3/P2′, and

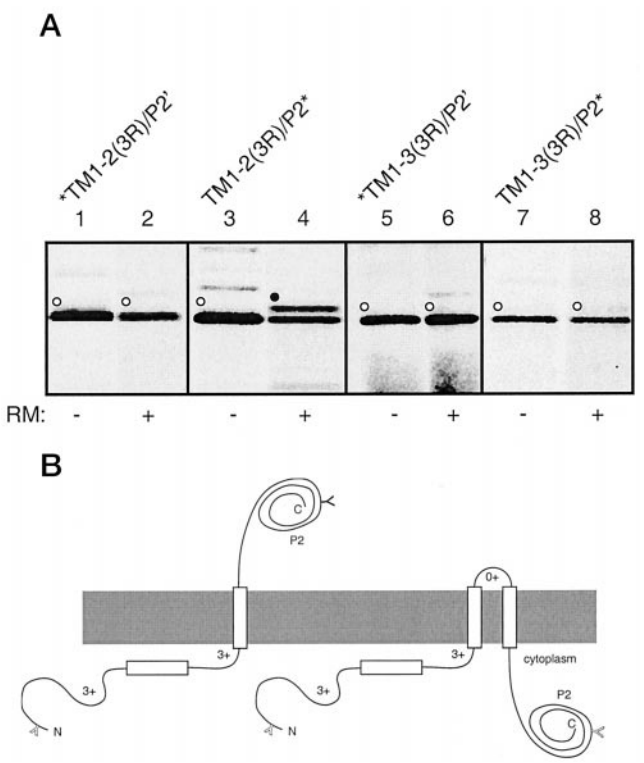

FIG. 4. **Addition of three Arg residues to the TM1–2 loop enhances its retention in the cytoplasm.** *A*, constructs *TM1–2(3R)/P2′ (*lanes 1* and *2*), TM1–2(3R)/P2* (*lanes 3* and *4*), *TM1–3(3R)/P2′ (*lanes 5* and *6*), and TM1–3(3R)/P2* (*lanes 7* and *8*) were expressed *in vitro* in the absence (−) and presence (+) of rough microsomes (*RM*). Nonglycosylated and glycosylated forms are indicated by *white* and *black dots*, respectively. *B*, topology models for the TM1–2(3R)/P2 (*left*) and TM1–3(3R)/P2 (*right*) constructs. *Solid Y-shaped symbol*, glycosylated site; *outlined Y-shaped symbol*, nonglycosylated site.

*TM1–3(3R)/P2′ fusions (Fig. 5*A*). In all cases, a unique Asn^3-Ser^4-Thr^5 glycosylation site was present in the N-tail. Little increase in N-tail glycosylation was seen for the shorter *TM1/P2′ and *TM1–2/P2′ constructs. In contrast, shortening of the N-tail in the *TM1–3/P2′ construct led to an appreciable increase in glycosylation efficiency. For the *TM1–3(3R)/P2′ construct, essentially complete N-tail translocation (*i.e.* ~80% modification) was observed when the N-tail was shorter than ~50 residues.

N-tail translocation also became markedly more efficient in the *TM1–3/P2′, *TM1–4/P2′, *TM1–5/P2′, *TM1–6/P2′, and *TM1–7/P2′ constructs when the N-tail was shortened to 30 residues (Fig. 5*B*; see also *black bars* in Fig. 3*C*), and similar results were obtained with the corresponding C-terminally truncated constructs (*white bars*; note that the *TM1(Δ1–70), *TM1–2(Δ1–70), and *TM1–3(Δ1–70) truncations were too poorly expressed to be analyzed).

Thus, N-tail translocation can be promoted by shortening the N-tail, and can be further increased by the addition of positively charged residues to the TM1–2 loop. Again, for a given length of the N-tail, we observe more efficient N-tail translocation as the number of TM segments is increased. Interestingly, the short N-tail in the *ProW(Δ1–70)/P2′-constructs is predominantly located in the lumen already when only three TM domains are present, in contrast to the full-length N-tail where four TM domains are required for significant translocation of the N-tail (Fig. 3*C*).

Efficient N-tail Translocation by Replacement of the ProW TM Domain—Because the efficiency of N-tail translocation was found to depend mainly on the number of transmembrane segments, it was also of interest to check whether the identity

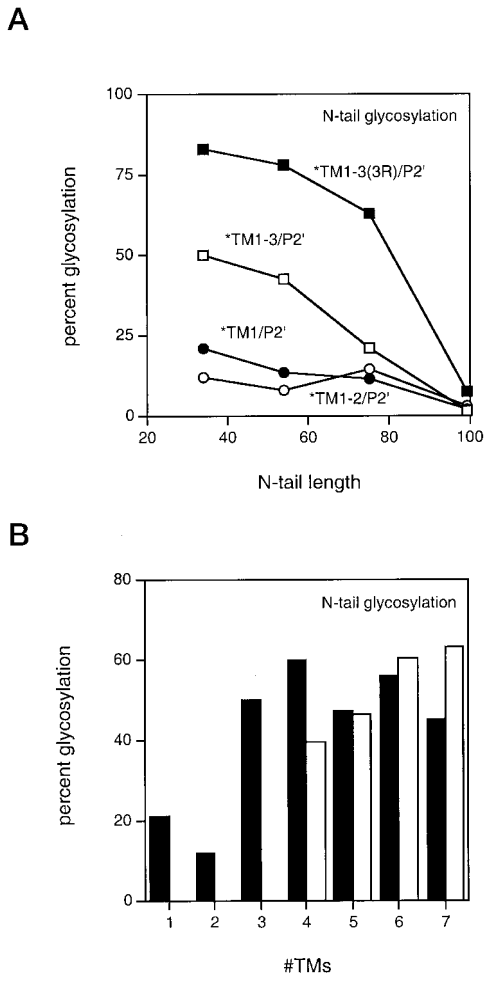

FIG. 5. **Shorter N-tails are more efficiently translocated**. *A*, degree of glycosylation of the N-tail in *TM1/P2′ (*black circles*), *TM1–2/P2′ (*white circles*), *TM1–3/P2′ (*white squares*), and *TM1–3(3R)/P2′ (*black squares*) as a function of the number of residues in the N-tail. *B*, degree of glycosylation of the N-tail in *ProW(Δ1–70)/P2′ constructs (*black bars*) and C-terminally truncated *ProW constructs (*white bars*).

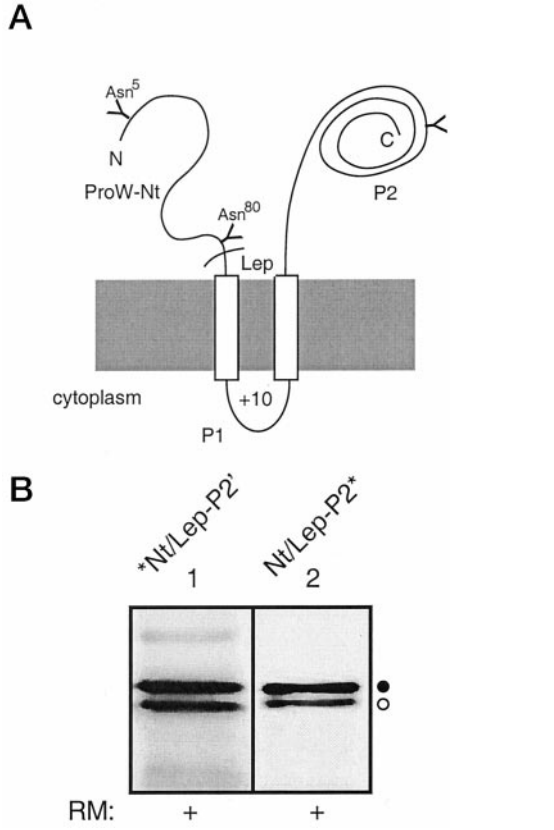

FIG. 6. **Efficient translocation of the ProW N-tail when fused to the N terminus of Lep**. *A*, the Nt/Lep fusion proteins (*solid Y-shaped symbol*, glycosylated site). *B*, the *Nt/Lep-P2′ (*lane 1*) and Nt/Lep-P2* (*lane 2*) fusions were expressed *in vitro* in the presence of rough microsomes (*RM*). Nonglycosylated and glycosylated forms are indicated by *white* and *black dots*, respectively.

of the transmembrane domain would affect the translocation efficiency. We thus constructed fusions between the ProW N-tail (residues 1–99) and Lep residues 1–215 (*Nt/Lep-P2′) or the full-length Lep protein (Nt/Lep-P2*). Lep has two transmembrane segments (residues 4–22 and 62–76), and both the N and C terminus face the lumen when Lep is integrated into microsomes (31, 32) (Fig. 6*A*). As shown in Fig. 6*B*, the ProW N-tail was quite efficiently translocated in this context (59% glycosylation of *Nt/Lep-P2′), despite the fact that only two TM segments are present, and the P2 domain was likewise translocated into the lumen as seen from the efficient glycosylation (62%) of Nt/Lep-P2*.

Kinetics of N-tail Glycosylation in Nt/Lep-P2′ Suggests That the N-tail Is Translocated in a C-to-N-terminal Direction—The rather efficient translocation of the full-length ProW N-tail in the *Nt/Lep-P2′ construct made it possible to study the direction of N-tail translocation. To this end, glycosylation acceptor sites were introduced in position 5 or 80 near the N-terminal and C-terminal ends of the N-tail, respectively (Fig. 6*A*). Synchronized translation was initiated by the addition of mRNA to the translation mix, and further chain initiation was blocked after 1.5 min by addition of the inhibitor aurintricarboxylic acid. The kinetics of glycosylation of the two sites was then followed as initially described by Rothman and Lodish (33), *i.e.* by dissolving the microsomes by addition of the detergent Triton X-100 at different time points (thus preventing further glycosylation) and then allowing translation to go to completion. Results for the *Nt(Asn^5)/Lep-P2′ and *Nt(Asn^{80})/Lep-P2′ constructs are shown in Fig. 7*A*. To rule out that the glycosylated forms of the two constructs represent different populations of molecules, the analysis was carried out also for construct **Nt(Asn^5, Asn^{80})/Lep-P2′ containing both sites.

As is clear from the quantitations in Fig. 7*B*, glycosylation of the *Nt(Asn^{80})/Lep-P2′ construct preceded that of the *Nt(Asn^5)/Lep-P2′ construct by about 3 min. Glycosylation of the Asn^{80} site was first detected about 7 min after initiation of translation. Separate measurement of the overall translation rate of the two constructs in the *in vitro* system yielded a value of 0.4 amino acids/s (data not shown). Assuming a uniform translation rate, the nascent chain is thus roughly 170 residues long after 7 min. Because approximately 70 residues are required to span the distance between the ribosomal P-site and the oligosaccharyl transferase active site (34), glycosylation of Asn^{80} happens at about the time expected if translocation is initiated by TM1, whereas Asn^5 is glycosylated ~3 min later. Glycosylation of the **Nt(Asn^5, Asn^{80})/Lep-P2′ construct proceeded in two steps separated by about 3 min, and the final glycosylation levels were 53% doubly glycosylated molecules and 14% singly glycosylated molecules (data not shown), consistent with the results for the single acceptor site mutants.

These results strongly suggest that the N-tail is translocated into the lumen in a C-to-N-terminal direction, starting with the insertion of TM1 into the translocation apparatus. Interestingly, the rate of N-tail translocation estimated from this experiment is quite slow and roughly of the same order as the rate of chain elongation.

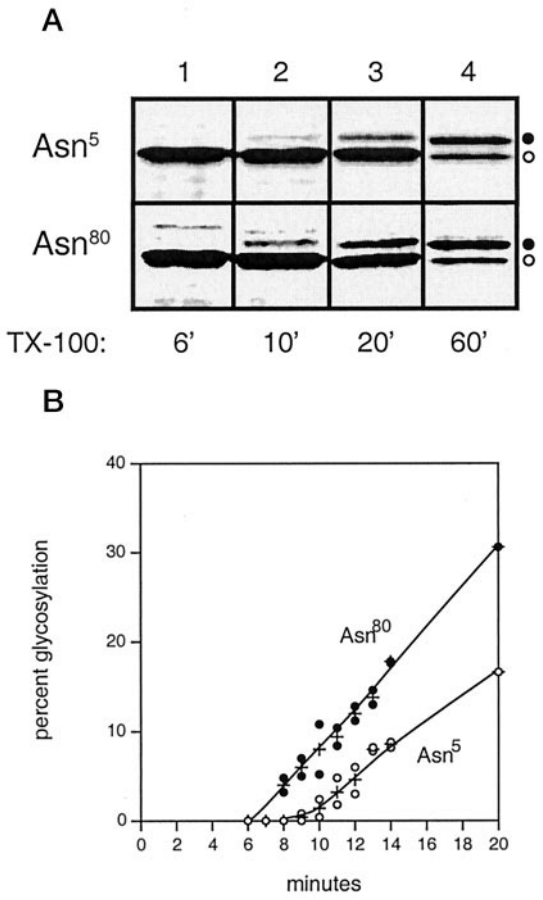

FIG. 7. **The ProW N-tail is translocated in a C-to-N-terminal direction**. The *Nt(Asn^5)/Lep-P2′ and *Nt(Asn^{80})/Lep-P2′ constructs were expressed *in vitro* in the presence of microsomes. After a 1.5-min incubation, aurintricarboxylic acid was added to block further initiation. Samples were removed at different time points, and Triton X-100 was added to dissolve the microsomes and block further glycosylation. Translation was then allowed to continue up to a total time of 60 min. *A*, selected time points of Triton X-100 addition for *Nt(Asn^5)/Lep-P2′ (*top*) and *Nt(Asn^{80})/Lep-P2′ (*bottom*). Nonglycosylated and glycosylated forms are indicated by *white* and *black dots*, respectively. *B*, kinetics of glycosylation determined from two independent experiments for each construct. *White dots*, *Nt(Asn^5)/Lep-P2′; *black dots*, *Nt(Asn^{80})/Lep-P2′). Mean values are shown by *crosses*.

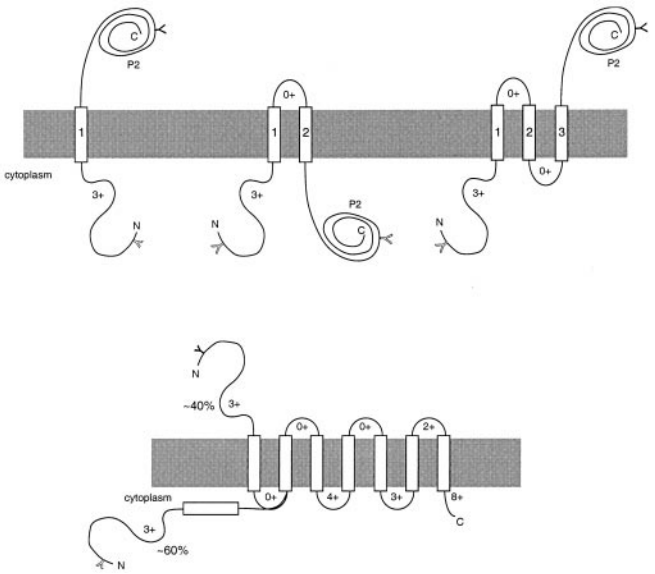

FIG. 8. **Topology models for the TM1/P2, TM1–2/P2, and TM1–3/P2 fusions (*top*) and for full-length ProW (*bottom*) inserted into microsomal membranes.** *Solid Y-shaped symbol*, glycosylated site; *outlined Y-shaped symbol*, nonglycosylated site. The number of positively charged residues (Arg+Lys) in the different loops are shown.

DISCUSSION

In this paper, we have analyzed the insertion of the *E. coli* inner membrane protein ProW into ER-derived microsomal membranes. In *E. coli,* the 100-residues long ProW N-tail is efficiently translocated across the inner membrane in the TM1–3/P2 construct, whereas in the shorter TM1/P2 construct it is translocated in about 50% of the molecules (6). The lack of positively charged residues immediately downstream of TM1 has been shown to be at least partly responsible for the appearance of TM1/P2 molecules with an inverted N_{cyt}-C_{exo} topology (7, 13).

The most unexpected finding in the present work is that very little N-tail translocation across the microsomal membrane is seen unless a minimum of four TM segments from ProW are present (Fig. 8). Instead, an inverted topology with the N-tail on the cytoplasmic side of the membrane dominates for the TM1, TM1–2, and TM1–3 constructs as determined by the glycosylation status of the N-tail and two different C-terminal reporter domains. Only when a minimum of four TM segments are present do we find a significant fraction of molecules with the N-tail in the lumen. It thus appears that the whole TM1-TM4 region has a direct influence on the translocation of the N-tail. In general, the length of the C-terminal tail does not influence N-tail translocation, except for the TM1–4 constructs where the N-tail is more efficiently translocated in the *TM1–4/P2′ fusion than when the protein is truncated after TM4. Because in the truncated *TM1–4 construct TM4 is still inside the ribosome when chain termination happens, it may not be able to influence the overall topology of the molecule in this case.

Insertion of three positively charged residues in the TM1-TM2 loop partially induces a topology with the TM1-TM2 loop in the cytoplasm and the TM2-TM3 loop in the lumen but does not lead to translocation of the N-tail unless it is shortened to 75 residues or less (Fig. 5*A*). Even without the extra positively charged residues in the TM1–2 loop, shortening the N-tail to ~30 residues leads to efficient N-tail translocation for all constructs with three or more TM domains. Finally, when the ProW N-tail is fused to another *E. coli* inner membrane protein (Lep) with only two TM segments, it is quite efficiently translocated.

Taken together, these results show that the balance between the N_{cyt} and N_{exo} orientations of the N-tail can be affected by sequence determinants more than 100 residues downstream of the tail itself (*i.e.* TM4 in ProW). This suggests that the final "decision" of whether or not to translocate the N-tail can be influenced by the topological preferences of TM segments that have not even been synthesized when TM1 enters the ER translocon, because only ~40 residues are required to span the distance between the translocon and the ribosomal P-site (35). Two extreme models for how this could happen are that the more N-terminal TM segments initially integrate into the ER membrane with the N-tail in the cytoplasm and are then pulled back into the translocon where they reorient as the more C-terminal TMs appear or that the whole protein remains in an "undecided" state inside the translocon until at least four TM segments have been made. At present, we cannot say which of the two models is closer to the truth, because experimental data in favor of both have been reported in the literature (4, 36–40). It is of course also possible or even likely that different proteins make the "topological decision" at different times during their biosynthesis, as suggested by the rather efficient

N-tail translocation observed for the *Nt/Lep-P2′ construct and for the ProW constructs with shortened N-tails.

Finally, analysis of the glycosylation kinetics of two *Nt/Lep-P2′ constructs with glycosylation sites placed either N- or C-terminally in the N-tail has allowed us to address a long standing question, namely whether N-tail translocation proceeds in an N-to-C-terminal or C-to-N-terminal direction. The results strongly suggest the latter and further imply that the rate of N-tail translocation is of the same order as the rate of chain elongation, at least in our *in vitro* system.

Acknowledgment—Dog pancreas microsomes were a kind gift from Dr. M. Sakaguchi.

REFERENCES

1. Blobel, G. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 1496–1500
2. Gafvelin, G., Sakaguchi, M., Andersson, H., and von Heijne, G. (1997) *J. Biol. Chem.* **272,** 6119–6127
3. Zhang, J.-T., Chen, M., Han, E., and Wang, C. (1998) *Mol. Biol. Cell* **9,** 853–863
4. Hegde, R. S., and Lingappa, V. R. (1999) *Trends Cell Biol.* **9,** 132–137
5. Dalbey, R. D., Kuhn, A., and von Heijne, G. (1995) *Trends Cell Biol.* **5,** 380–383
6. Whitley, P., Zander, T., Ehrmann, M., Haardt, M., Bremer, E., and von Heijne, G. (1994) *EMBO J.* **13,** 4653–4661
7. Whitley, P., Gafvelin, G., and von Heijne, G. (1995) *J. Biol. Chem.* **270,** 29831–29835
8. Cao, G. Q., and Dalbey, R. E. (1994) *EMBO J.* **13,** 4662–4669
9. Mitsopoulos, C., Hashemzadeh-Bonehi, L., and Broome-Smith, J. (1997) *FEBS Lett* **419,** 18–22
10. McMurry, J. L., and Kendall, D. A. (1999) *J. Biol. Chem.* **274,** 6776–6782
11. Oliver, J., Jungnickel, B., Görlich, D., Rapoport, T., and High, S. (1995) *FEBS Lett.* **362,** 126–130
12. Denzer, A. J., Nabholz, C. E., and Spiess, M. (1995) *EMBO J.* **14,** 6311–6317
13. Cristobal, S., Scotti, P., Luirink, J., von Heijne, G., and de Gier, J. W. L. (1999) *J. Biol. Chem.* **274,** 20068–20070
14. Wallin, E., and von Heijne, G. (1995) *Protein Eng.* **8,** 693–698
15. Delgado-Partin, V. M., and Dalbey, R. E. (1998) *J. Biol. Chem.* **273,** 9927–9934
16. Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 5786–57890
17. Monné, M., Gafvelin, G., Nilsson, R., and von Heijne, G. (1999) *Eur. J. Biochem.* **263,** 264–269
18. Haardt, M., and Bremer, E. (1996) *J. Bacteriol.* **178,** 5370–5381
19. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 488–492
20. Geisselsoder, J., Witney, F., and Yuckenberg, P. (1987) *BioTechniques* **5,** 786–791
21. Johnston, S., Lee, J. H., and Ray, D. S. (1985) *Gene* (*Amst.*) **34,** 137–145
22. Kozak, M. (1992) *Annu. Rev. Cell Biol.* **8,** 197–225
23. Whitley, P., Nilsson, I., and von Heijne, G. (1994) *Nat. Struct. Biol.* **1,** 858–862
24. Liljeström, P., and Garoff, H. (1991) *J. Virol.* **65,** 147–154
25. Sakaguchi, M., Mihara, K., and Sato, R. (1987) *EMBO J.* **6,** 2425–2431
26. Lyman, S., and Schekman, R. (1997) *Cell* **88,** 85–96
27. van Geest, M., Nilsson, I., von Heijne, G., and Lolkema, J. S. (1999) *J. Biol. Chem.* **274,** 2816–2823
28. Nakai, T., Yamasaki, A., Sakaguchi, M., Kosaka, K., Mihara, K., Amaya, Y., and Miura, S. (1999) *J. Biol. Chem.* **274,** 23647–23658
29. Shakin-Eshleman, S. H., Wunner, W. H., and Spitalnik, S. L. (1993) *Biochemistry* **32,** 9465–9472
30. Fujiki, Y., Hubbard, A. L., Fowler, S., and Lazarow, P. B. (1982) *J. Cell Biol.* **93,** 97–102
31. Johansson, M., Nilsson, I., and von Heijne, G. (1993) *Mol. Gen. Genet.* **239,** 251–256
32. Nilsson, I., and von Heijne, G. (1993) *J. Biol. Chem.* **268,** 5798–5801
33. Rothman, J., and Lodish, H. (1977) *Nature* **269,** 775–780
34. Whitley, P., Nilsson, I. M., and von Heijne, G. (1996) *J. Biol. Chem.* **271,** 6241–6244
35. Mothes, W., Prehn, S., and Rapoport, T. A. (1994) *EMBO J.* **13,** 3937–3982
36. Mothes, W., Heinrich, S., Graf, R., Nilsson, I., von Heijne, G., Brunner, J., and Rapoport, T. (1997) *Cell* **89,** 523–533
37. Do, H., Falcone, D., Lin, J., Andrews, D. W., and Johnson, A. E. (1996) *Cell* **85,** 369–378
38. Borel, A. C., and Simon, S. M. (1996) *Cell* **85,** 379–389
39. Ota, K., Sakaguchi, M., von Heijne, G., Hamasaki, N., and Mihara, K. (1998) *Mol. Cell* **2,** 495–503
40. Goder, V., Bieri, C., and Spiess, M. (1999) *J. Cell Biol.* **147,** 257–266