Signal Sequences Initiate the Pathway of Maturation in the Endoplasmic Reticulum Lumen*

Received for publication, February 28, 2003, and in revised form, April 21, 2003
Published, JBC Papers in Press, May 27, 2003, DOI 10.1074/jbc.M302117200

D. Thomas Rutkowski‡‡, Carolyn M. Ott††, Jon R. Polansky‡, and Vishwanath R. Lingappa**‡‡

From the ‡Department of Biochemistry and Biophysics and the **Departments of Physiology and Medicine and ††Ophthalmology, University of California San Francisco, San Francisco, California 94143

An interaction between an N-terminal signal sequence and the translocon leads to the initiation of protein translocation into the endoplasmic reticulum lumen. Subsequently, folding and modification of the substrate rapidly ensue. The close temporal coordination of these processes suggests that they may be structurally and functionally coordinated as well. Here we show that information encoded in the hydrophobic domain of a signal sequence influences the timing and efficiency of at least two steps in maturation, namely N-linked glycosylation and signal sequence cleavage. We demonstrate that these consequences correlate with and likely stem from the nature of the initial association made between the signal sequence and the translocon during the initiation of translocation. We propose a model by which these maturation signals are controlled by the signal sequence-translocon interaction. Our work demonstrates that the pathway taken by a nascent chain through post-translational maturation depends on information encoded in its signal sequence.

N-terminal signal sequences enable nascent secretory and transmembrane proteins to be targeted to the endoplasmic reticulum (ER) for translocation. The signal sequence, newly emerged from the translating ribosome, is recognized in the cytoplasm by the signal recognition particle (1). By virtue of an interaction with its ER-localized receptor, signal recognition particle brings the ribosome-nascent chain complex to the ER membrane (2). Once at the ER, the ribosome is thought to facilitate the assembly or stabilization of the translocation channel, composed of the heterotrimeric Sec61 complex (3, 4). The 35-kDa polytopic Sec61α subunit appears to form the channel walls, whereas the smaller bitopic β- and γ-subunits play an as yet undetermined role, perhaps in facilitating the insertion of the nascent chain into the channel (5–7).

An important advance in the understanding of the initiation of translocation was the discovery of a second step of signal sequence recognition. In addition to being recognized by the signal recognition particle, the signal sequence also interacts with the translocation channel itself (8, 9). Although the mechanism is not yet clear, this event is thought to stimulate the initiation of translocation. For the simplest secretory proteins, when the signal sequence is recognized by the channel, the ribosome assumes a tight interaction with the channel that shields the protein from the cytoplasm (8, 10). Concomitantly, the channel opens toward the ER lumen, either via a conformational change in the channel or removal of a molecular plug covering the luminal aperture (11, 12).

Not all signal sequences initiate translocation in the same way. Beyond merely stimulating the initiation of translocation, signal sequences can regulate the association between the ribosome and translocon and the exposure of the nascent chain to the cytoplasm or ER lumen (13–16). In at least one case, that of the prion protein, the consequence of this regulatory step is the governance of the final topology achieved by the protein (13–17). These observations raise the possibility that signal sequences have a broad post-targeting role beyond the initiation of translocation, and in addition to their role in regulating the ribosome-translocon association.

Immediately after translocation begins, the protein complexes of the ER lumen begin the process of maturation of the nascent chain. Maturation includes steps of covalent modification such as signal sequence cleavage, the addition of N-linked glycans, and glycosylphosphatidylinositol linkage as well as protein folding and refolding carried out by chaperones in the lumen (18). It is clear that the pathway of a nascent chain through the steps of maturation in the ER lumen varies from substrate to substrate (19). However, the manner in which the steps of biogenesis are optimized for a particular protein is not yet clear.

Given that signal sequence recognition is the first event in translocation at the ER and also that maturation is closely linked to translocation, we wondered whether the signal sequence-translocon interaction might control the progression of the nascent chain through the steps of maturation in the ER lumen. We demonstrate that signal sequences can indeed influence the timing and efficiency of maturational events, and this influence can be conferred to other proteins on which those signal sequences are engineered. This influence is likely exerted through the nature of the interaction between the signal sequence and translocon.

EXPERIMENTAL PROCEDURES

Plasmid Constructions—cDNAs encoding substrates synthesized in the cell-free system were cloned into the pSP64 vector (Promega), and cDNAs encoding transfected substrates were cloned into pcDNA 3.1 Neo (+) (Invitrogen). The glycosylation site in Prl was made by site-directed mutagenesis of Gln to Asn using the QuikChange kit (Stratagene). Prl constructs with various signal sequences (Fig. 2A) were synthesized by PCR amplification of these signal sequences and ligation into an

---

* 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.

† A former Howard Hughes Predoctoral Fellow in Biological Sciences.

‡‡ Supported by grants from the National Institutes of Health and Sandler Foundation. To whom correspondence should be addressed. Tel.: 415-476-2762; Fax: 415-476-4929; E-mail: vr3@itsa.ucsf.edu

1 The abbreviations used are: ER, endoplasmic reticulum; RM, canine pancreatic rough microsomes; PDI, protein-disulfide isomerase; Prl, preprolactin; TIGRp, trabecular meshwork inducible glucocorticoid response protein; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; GH, growth hormone; AP, acceptor tripeptide.
upstream Sphl site and a silent AgeI site located in the mature domain of Prl. The plasmids encoding growth hormone (GH) and IgG were provided by Tom Rapoport (Harvard Medical School, Boston, MA) and Prl receptor by Linda Schuler (University of Wisconsin, Madison, WI). Chimeras of PrlCHO and GH-PrlCHO (Fig. 2B) were made by subcloning into upstream Sphl, NcoI, or Hind3 sites, a Pmsp MI site between the Prl α- and h-domains, a Pfml site between the h- and c-domains, and the silent HinI site. Libraries of Prl hydrophobic domain mutations were constructed by ligation of degenerate oligonucleotides of the sequence (NYEQ5)12 into the Pmsp MI and Pfml sites of PrlCHO (Fig. 2C), or by random QuikChange mutagenesis of Ser26-Asn34. Unchargeable signal sequences (Fig. 5D) were generated by mutation of GVV5 of the PrlCHO or SN or —Q5CN6, a signal sequence to WPVP. Standard techniques were used in the construction of all clones (20).

In Vitro Translation, Translocation, and Proteolysis—In vivo translation, translocation, and proteolysis have been described previously (10). Where indicated, substrates were synthesized in the presence of a tripeptide inhibitor of glycosylation, Ac-MTyr. Truncated substrates were generated by PCR amplification of fragments of the relevant cDNAs, followed by transcription and translation. For targeting analysis (Fig. 4A), chains truncated at the indicated lengths were synthesized in the presence of RMs. Half of the material was set aside, whereas half was sedimented for 3 min at 50,000 rpm and 4 °C in a TLA 100 rotor over a 50-μl cushion of 0.5 m sucrose, 0.5 m KOAc, 50 mM Hepes, pH 7.5, and 5 mM MgOAc. Both total material and sedimented material were precipitated with 50% ammonium sulfate at 0 °C for 30 min. Analysis of the state of the ribosome-translocon junction (Fig. 4B) was exactly as described (16). For puromycin analysis (Fig. 6B), isolated RMs with associated targeted chains were resuspended in 250 mM KOAc, 50 mM Hepes (pH 7.5), 5 mM MgOAc, and 0.25 m sucrose with or without 1 m puromycin (Sigma). Reactions were incubated at 32 °C for 15 min, followed by proteolysis.

Tryptic Analysis—For tryptic digestion, Prl truncated at 80 amino acids was synthesized in the presence or absence of RMs at 32 °C for 20 min. Samples synthesized in the absence of RMs were isolated by sedimentation of the polysomes for 30 min at 100,000 rpm, 4 °C, in a TLA 100 rotor over a 100-μl cushion of 0.5 m sucrose, 100 mM KOAc, 50 mM Hepes, pH 7.5, and 5 mM MgOAc. Both polysomes and sedimented RMs were resuspended in 250 μl of the same sucrose cushion. Both polysomes and sedimented RMs were resuspended in 0.25 m sucrose, 100 mM KOAc, 50 mM Hepes, pH 7.5, and 5 mM MgOAc plus 1% Triton X-100. Digestion was carried out with 10 μg trypsin (Sigma) for 10 min at 0 °C. Trypsin was inactivated with 5 mM phenylmethylsulfonyl fluoride for 5 min at 0 °C, followed by transfer to tubes containing boiling 1% SDS. Samples were separated by SDS-PAGE and exposed to autoradiography. The films were scanned, and the intensities of bands in each lane were quantitated by measuring the densities of small increments of each lane in progressive intervals, from high molecular weight to low. Thus, the y axis of the graphs in Fig. 6A is in arbitrary units.

Cross-linking and Immunoprecipitation—In Fig. 7, following translation, targeted chains were recovered by sedimentation over a sucrose cushion and resuspended in a physiological salt buffer (10) (as before tryptic digestion, but without 1% detergent). Half of the material was set aside, and half was incubated for 80 min at room temperature with a 1 mM concentration of the cysteine-reactive homobifunctional reagent 1,8-bis-maleimidothiethylenglycol (Pierce), from a 20 mM stock dissolved in water. Just prior to its addition to the reactions, cross-linker was incubated with agitation at 40 °C to stimulate dissolution. The reactions were quenched with 1 volume of 20 mM dithiothreitol, 0.1% Triton X-100, 15 mM EDTA, and 20 mM PMSF and 10% glycoproteinase A at room temperature for 10 min. For immunoprecipitation with anti-PDI (Stressgen), 0.5% saponin was added to the quenched reaction, and the samples were incubated with antibody overnight at 4 °C before dilution into 1× TXSWB (1% Triton X-100, 0.1 mM Hepes, pH 7.5, 0.1 mM NaCl) and the addition of Protein A-agarose. A nonimmune control immunoprecipitation experiment was performed similarly (not shown). For other immunoprecipitations, quenched reactions were denatured with the addition of 0.5% SDS and heating to 100 °C. The samples were then diluted in 10 volumes of 1× TXSWB and incubated with antibody and Protein A-agarose overnight or with anti-Sec61α coupled to Sepharose provided by Tom Rapoport.

Cell Culture—COS-1 cells were cultured in high glucose Dulbecco’s modified Eagle’s medium (University of California San Francisco Cell Culture Facility) with 10% fetal bovine serum and antibiotics. In Fig. 1C, lysates or media were harvested, denatured in SDS, and immunoblotted with a polyclonal antibody raised against TIGRp.

**RESULTS**

In order to test our hypothesis that signal sequences influence the progression of nascent chains through maturation in the ER lumen, we monitored the glycosylation of the trabecular meshwork inducible glucocorticoid response protein (TIGRp), also known as myocilin (23). TIGRp is a secreted protein and contains a single site for N-linked glycosylation, at position 57 of the nascent chain, which is 25 amino acids downstream of the site of signal sequence cleavage (Fig. 1A). We replaced the signal sequence of TIGRp with that of the secretory protein preprolactin (Prl) and compared the extents to which native TIGRp and TIGRp with the Prl signal sequence (PrlTIGRp) were glycosylated. We found that PrlTIGRp was glycosylated with a significantly lower efficiency, both in vitro and in vivo, than was TIGRp, despite the fact that both substrates translate with equally high efficiency (Fig. 1, B and C, and data not shown). Thus, the efficiency of at least one step (i.e. N-linked glycosylation) in the maturation of this native substrate is affected by the signal sequence of the nascent chain in a manner distinct from its ability to be translocated per se.

We chose Prl as a substrate for further analysis because the steps in its translocation have been thoroughly characterized.
Thus, we reasoned that this substrate would best allow us to define the mechanism by which the signal sequence, which functions in translocation, can influence maturation. We engineered into the Prl mature domain, by point mutagenesis, a single site for glycosylation at amino acid number 42, which is 12 amino acids downstream of the site of signal sequence cleavage. The glycosylation of this substrate (PrlCHO) was compared with glycosylation of the PrlCHO mature domain fused to the signal sequences of various other proteins, including the prolactin receptor, GH, and IgG heavy chain. As with TIGRp, PrlCHO glycosylation was influenced to a large extent by the signal sequence used to direct its translocation; whereas the Prl receptor and native Prl signal sequences led to inefficient glycosylation of translocated chains, the GH and IgG heavy chain signal sequences allowed a majority of translocated chains to be modified (Fig. 2A). Thus, we confirmed that Prl biogenesis could be used as a reporter for the effects of signal sequences on maturation in the ER lumen.

Signal sequences contain a central hydrophobic core of 6–15 amino acids, flanked by N-terminal and C-terminal polar domains of varying length. The hydrophobic domain is required for targeting and interaction with the translocon (1, 8). The N-terminal domain is thought to determine the dependence of the substrate on the presence of the translocating chain-associated membrane protein (27), and the c-terminal domain governs the site of signal sequence cleavage (28). To determine which of these domains is responsible for affecting glycosylation, PrlCHO is shown by the open diamond.
Signal Sequences Influence ER Maturation

Common properties of the signal sequence underlie glycosylation of PrlCHO and TIGRp. A, Sec20 and Asn21 of the Prl signal sequence were mutated to generate the substrates indicated. The arrowhead represents the site of signal sequence cleavage. In general, when glycosylation-competent clones are used (for which Gln42 is changed to Asn), they are designated as CHO. The asterisks show the locations of cysteine residues reactive to a thiol-specific homobifunctional cross-linking agent (see Fig. 7). B, PrlCHO, SN → NHCHO, SN → MPCHO, and SN → QTCHO were synthesized, and the extent of glycosylation was quantitated as in Fig. 1B. C, the signal sequences shown in A were fused to the mature domain of TIGRp, and glycosylation was monitored as usual.

content of the hydrophobic domain used (Fig. 2D). Therefore, information in the hydrophobic domain that regulates glycosylation efficiency is not encoded solely by any one of these parameters and perhaps is encoded in some other manner.

To best pinpoint the mechanism by which a signal sequence could influence glycosylation, we wanted to isolate discrete mutations in the Prl signal sequence that could lead to changes in glycosylation while leaving functional properties related to translocation as intact as possible. To accomplish this goal, we randomly mutated Sec20 and Asn21 of the Prl signal sequence, since these polar residues are probably not required for targeting or the initiation of translocation, both of which are thought to rely on hydrophobic residues within the signal sequence (1, 8). The substrates generated by this technique varied widely in glycosylation efficiency, so for further study we chose one signal sequence (SN → NH) that, like the native Prl signal, led to inefficient modification and two (SN → MP and SN → QT) that led to very efficient glycosylation (Fig. 3A). The glycosylation of SN → MPCHO and SN → QTCHO was ~3-fold more efficient than for PrlCHO and SN → NHCHO (Fig. 3B). Importantly, these signal sequences had similar effects when used to direct TIGRp translocation (Fig. 3C), suggesting that glycosylation of both TIGRp, a native glycoprotein, and Prl, an engineered glycoprotein, are governed by the same functional properties of the signal sequence.

We next turned our attention toward the properties of these different signal sequences that lead to differences in glycosylation. In understanding how these signal sequences affect maturation, it was important to identify ways in which they differed independent of the potential of the chain to be glycosylated and then determine if any such differences predicted the final efficiency of glycosylation that could be achieved. Thus, we created versions of PrlCHO, SN → QTCHO, SN → MPCHO, and SN → NHCHO that lacked the engineered glycosylation site (i.e. these substrates were composed of the native Prl mature domain fused to the different signal sequences and are named Prl, SN → QT, SN → MP, and SN → NH). In all subsequent experiments when glycosylation itself (e.g. its timing or efficiency) was not considered, these glycosylation-competent substrates were used (Figs. 4, 5, A and B, 6, and 7). It is worth noting, however, that the glycosylation-competent versions of these substrates displayed indistinguishable behavior compared with their incompetent counterparts in all cases (data not shown).

The best characterized role of signal sequences is in mediating targeting to the ER. Whereas the fact that the PrlCHO mature domain is translocated with high efficiency irrespective of the signal sequence used makes it unlikely that glycosylation is impacted by the targeting property of the signal sequence, for mechanistic insight it was important to formally exclude this possibility. We synthesized Prl and SN → QT nascent chains, truncated at increasing lengths, in the presence of ER-derived microsomal membranes. Targeted chains were then isolated by sedimenting the microsomes through a cushion of sucrose and high salt, and the percentage of material recovered in this way was quantitated. This analysis revealed that the timing of targeting for Prl and SN → QT is indistinguishable (Fig. 4A).

The best currently described posttargeting role of signal sequences in maturation is in influencing the ribosome-translocon interaction (14–16). To determine if signal sequences could influence glycosylation independent of their role in regulating this association, we examined the timing with which Prl and SN → QT achieved a tight ribosome-translocon association. To accomplish this, we determined the percentage of nascent chains, truncated at various lengths, that were protected from digestion by exogenously added protease K, which has access only to material exposed to the cytoplasm. Both Prl and SN → QT achieved protease-resistant binding to the ER membrane with the same timing (Fig. 4B). Therefore, it seems likely that additional levels of posttargeting regulation besides that of the ribosome-translocon junction impinge upon...
the nascent chain. We proceeded to explore the mechanism by which signal sequences exert this influence.

Inhibiting signal sequence cleavage can in some cases prevent modification of potential glycosylation sites in translocation substrates (29). All of the substrates analyzed in this study undergo quantitative signal sequence cleavage (Fig. 2A and data not shown); therefore, differential glycosylation cannot be attributable to the failure of any of these signal sequences to be cleaved. However, signal sequences also vary with respect to the timing with which they are cleaved, although how such a property is encoded and recognized by the translocation machinery is not understood (28). We considered the possibility that the different signal sequences used to direct Prl-CHO translocation might vary in the timing of their own cleavage and that this property might be responsible for governing the efficiency of glycosylation. To test this hypothesis, we first asked whether the Prl and SN → QT signal sequences are in fact cleaved at different times. We synthesized translation intermediates of Prl and SN → QT, truncated at different points, and compared the extent of signal sequence cleavage achieved by these substrates at each point (Fig. 5A). This analysis revealed that the SN → QT signal sequence has undergone substantial cleavage by the synthesis of 136 amino acids of substrate, whereas a majority of Prl chains are not cleaved until 164 amino acids have been synthesized. To determine whether the timing of signal sequence cleavage is associated with glycosylation efficiency, we also monitored the timing of signal sequence cleavage for SN → MP and SN → NH (Fig. 5B). The SN → QT and SN → MP signal sequences, which lead to efficient glycosylation of the Prl-CHO mature domain, are cleaved earlier than the Prl and SN → NH signal sequences. Importantly, the timing of signal sequence cleavage did not depend on whether the substrate was glycosylation-incompetent (Fig. 5B) or glycosylation-competent (Fig. 5C and data not shown).

We next compared the timing of signal sequence cleavage and the timing of glycosylation for Prl-CHO and SN → QT-CHO. We found that for both substrates, signal sequence cleavage preceded glycosylation; no glycosylated but signal sequence-uncleaved material was present for either Prl-CHO or SN → QT-CHO (Fig. 5C). In addition, glycosylation followed rapidly after signal sequence cleavage; thus, substantial amounts of glycosylated material are seen as early as 136 amino acids for SN → QT but not until 164 amino acids for Prl (material marked by an asterisk in Fig. 5C). These data suggest that glycosylation cannot occur efficiently until the signal sequence has been cleaved. Indeed, when the Prl-CHO and SN → QT-CHO signal sequences were mutated to render them uncleavable, both substrates were glycosylated relatively inefficiently (Fig. 5D). Thus, it seems that whereas glycosylation can occur to a limited extent independent of whether or not the signal sequence has been cleaved, efficient glycosylation occurs during a relatively limited window of time in chain growth after signal sequence cleavage has occurred.

How can a signal sequence encode the timing of its own cleavage? The most plausible mechanism by which the timing of signal sequence cleavage could be differentially controlled in two signal sequences with identical cleavage sites is via the conformation or environment of the nascent chain. A nascent chain could conceivably be positioned in such a way with respect to the signal peptidase complex (and by extension, other proteins in and around the translocon) that the recognition site for the peptidase cannot easily access it. Altering elements within the signal sequence might change the positioning of the chain, increasing its accessibility to the peptidase. Thus, we wanted to determine whether the altered timing of signal sequence cleavage is indeed a reflection of the Prl and SN → QT signal sequences leading to alternate conformations for their associated nascent chains.

To further explore this possibility, we first used limited trypsin digestion, a classic probe of differences in protein conformation/association. We isolated short (80 amino acids) targeted chains of Prl and SN → QT substrates and solubilized the microsomal membranes in nondenaturing detergent before adding low concentrations of trypsin. When digestion was carried out on untargeted material (i.e. on isolated polyribosomes), both substrates yielded identical trypsin digestion patterns that consisted largely of a single species that probably represents the length of chain protected from digestion by the ribosome (Fig. 6A, −RM lanes). However, when the substrates were targeted to the ER, they showed differential digestion patterns (Fig. 6A, +RM lanes). Similar results were observed at later points in synthesis and also upon more or less aggressive digestion with trypsin (data not shown). From these trypsin digestion studies, we conclude that the Prl and SN → QT nascent chains are in different environments with respect to the translocon and associated proteins and/or lipids. Because the identical chains, when synthesized in the absence of ER, showed no such difference, we conclude that this effect was not a trivial consequence of the mutation per se. It is important to note that the arginine residues within both proteins are located far from the site of signal sequence cleavage (GenBank™ ac-
Signal Sequences Influence ER Maturation

In this study, we have followed the progression of a nascent chain through maturation in the ER lumen and found that maturation is controlled at least in part by information encoded in the chain’s signal sequence. In this case, this information controls the manner in which the nascent chain is presented to the machinery of maturation, which in turn affects the timing and/or efficiency of both signal sequence cleavage and glycosylation. These observations apply both to an engineered model

translocate with similar efficiency in their full-length forms. If Prl and SN → QT chains had been in identical conformations at the ER, then there would be no reason to presume that they would differ in their localization upon release by puromycin.

What is the nature of the differences in substrate environment at the level of protein-protein interactions, and is this effect responsible for changes in the timing of signal sequence cleavage? We used chemical cross-linking to probe the environment of the nascent chain at progressive intervals early during translocation, looking for differences in associations that precede and predict signal cleavage timing and glycosylation efficiency. Strikingly, when analyzing Prl and SN → QT, we found several differences in cross-linking (using the cysteine-reactive homobifunctional reagent 1,8-bis-maleimidotriethyleneglycol) that persisted over a significant length of the chains’ early biogenesis. Two bands, representing the Prl and SN → QT substrates cross-linked to proteins of 55 (p55) and 60 (p60) kDa (Fig. 7A), were seen differentially for the substrates between chain lengths of 95 and 125 amino acids. (The presence of these species provides additional evidence for Prl and SN → QT being in different conformations, although they do not correlate with glycosylation efficiency for other substrates (data not shown) and so were not identified). In addition, we also observed differences in the efficiency of cross-linking to species of 35 (p35) and 45 (p45) kDa (Fig. 7A). Specifically, the Prl substrate associated more strongly with both species than did SN → QT. This difference was observed even at a chain length of 70 amino acids (Fig. 7B), which is the point at which Prl first has access to the ER lumen (8, 12, 30). We also observed two cross-links (10 and 50 kDa) at this chain length that were seen more strongly for SN → QT than for Prl. Immunoprecipitation was used to identify the 50-kDa species as the luminal chaperone protein-disulfide isomerase (PDI), the 35-kDa species as Sec61α, and the 10-kDa species as Sec61β. The 45-kDa species, which is recognized by antibodies against both Sec61α and Sec61β, is Sec61β cross-linked to a dimer of the nascent chain and Sec61α (Sec61β has only a single cross-linker-reactive residue and therefore cannot form a covalent bridge between Sec61α and the nascent chain) (Fig. 7C).

Strikingly, we found that the association of the chain with Sec61α (as measured by the generation of the p45 trimeric complex versus Sec61β was predictive of signal sequence cleavage and glycosylation. SN → NH and Prl, which are inefficiently glycosylated, can be more readily cross-linked to Sec61α, whereas SN → MP and SN → QT, which are efficiently glycosylated, are more readily cross-linked to Sec61β (Fig. 7D).

We also found that, whereas the signal sequence leads to global alterations in the conformation of the nascent chain (Fig. 6), the signal sequence itself is in a different environment with respect to Sec61α. We come to this conclusion because the same differential cross-linking to Sec61α is observed when the Prl and SN → QT signal sequences are transplanted onto two additional mature domains (TIGRp and GHI), which lack additional cross-linker-reactive residues at the chain length used (Fig. 7E). This result demonstrates that the position of the signal sequence with respect to Sec61α is an intrinsic characteristic of the signal sequence, independent of input from the mature domain.

**FIG. 6.** Prl and SN → QT nascent chains are in distinct conformations. A. Prl and SN → QT truncated at a chain length of 80 amino acids were synthesized in the presence or absence of RM. Following synthesis, either polysomes (for − RM) or microsomes (for + RM) were isolated by sedimentation. The isolated material was resuspended in a physiological salt buffer containing 1% Triton X-100, which solubilizes microsomes but is nondenaturing. Material was either set aside, whereas the remainder was treated with 1 mM puromycin and 10 μM trypsin at 0 °C and subjected to SDS-PAGE. The bar graph to the right shows the percentage of material for each substrate that has achieved signal sequence cleavage after puromycin release.

To further explore our hypothesis that the Prl and SN → QT signal sequences lead to their associated nascent chains being presented in different ways to the machinery of maturation, we took an entirely different approach to probing the environment of the two populations of nascent chains. We used the aminoacyl-tRNA analog puromycin to release them from the ribosome and determined their final destination. Ordinarily, a chain that has initiated translocation will be released by puromycin into the ER lumen, where its presence is confirmed by its protection from digestion by exogenously added proteinase K. Using puromycin to release 80-mers of the Prl and SN → QT substrates, we found that more of the latter were released into the cytoplasm. This localization is evidenced by fewer SN → QT chains achieving signal sequence cleavage upon release and the majority of these uncleaved chains being susceptible to digestion by proteinase K in the cytoplasm (Fig. 6B). In contrast, for Prl both cleaved and uncleaved chains are well protected from digestion. The preferential release into the cytoplasm for SN → QT comes despite the fact that Prl and SN → QT target to the ER and achieve stable protease-resistant binding to the translocation channel with the same timing (Figs. 4, A and B) and

cession number V00112). These data, combined with the effects of these signal sequences on the timing of cleavage, suggest that the effects of the signal sequence on nascent chain conformation or environment are not limited strictly to the site of signal sequence cleavage but extend over a large portion of the length of the nascent chain.

To further explore our hypothesis that the Prl and SN → QT signal sequences lead to their associated nascent chains being presented in different ways to the machinery of maturation, we took an entirely different approach to probing the environment of the two populations of nascent chains. We used the aminoacyl-tRNA analog puromycin to release them from the ribosome and determined their final destination. Ordinarily, a chain that has initiated translocation will be released by puromycin into the ER lumen, where its presence is confirmed by its protection from digestion by exogenously added proteinase K. Using puromycin to release 80-mers of the Prl and SN → QT substrates, we found that more of the latter were released into the cytoplasm. This localization is evidenced by fewer SN → QT chains achieving signal sequence cleavage upon release and the majority of these uncleaved chains being susceptible to digestion by proteinase K in the cytoplasm (Fig. 6B). In contrast, for Prl both cleaved and uncleaved chains are well protected from digestion. The preferential release into the cytoplasm for SN → QT comes despite the fact that Prl and SN → QT target to the ER and achieve stable protease-resistant binding to the translocation channel with the same timing (Figs. 4, A and B) and

cession number V00112). These data, combined with the effects of these signal sequences on the timing of cleavage, suggest that the effects of the signal sequence on nascent chain conformation or environment are not limited strictly to the site of signal sequence cleavage but extend over a large portion of the length of the nascent chain.

To further explore our hypothesis that the Prl and SN → QT signal sequences lead to their associated nascent chains being presented in different ways to the machinery of maturation, we took an entirely different approach to probing the environment of the two populations of nascent chains. We used the aminoacyl-tRNA analog puromycin to release them from the ribosome and determined their final destination. Ordinarily, a chain that has initiated translocation will be released by puromycin into the ER lumen, where its presence is confirmed by its protection from digestion by exogenously added proteinase K. Using puromycin to release 80-mers of the Prl and SN → QT substrates, we found that more of the latter were released into the cytoplasm. This localization is evidenced by fewer SN → QT chains achieving signal sequence cleavage upon release and the majority of these uncleaved chains being susceptible to digestion by proteinase K in the cytoplasm (Fig. 6B). In contrast, for Prl both cleaved and uncleaved chains are well protected from digestion. The preferential release into the cytoplasm for SN → QT comes despite the fact that Prl and SN → QT target to the ER and achieve stable protease-resistant binding to the translocation channel with the same timing (Figs. 4, A and B) and
glycoprotein (Prl) and to a native glycoprotein (TIGRp), suggesting their generality.

The requirement for an interaction between the signal sequence and translocation channel to stimulate the initiation of translocation is well established (8, 9). However, our data suggest that there is not a single universal mode of interaction between a signal sequence and the Sec61 complex but instead that a signal sequence may interact with components of the translocon in multiple distinguishable states. Each of these is apparently capable of stimulating the initiation of translocation, yet each has particular consequences on the maturational events that occur once translocation has been initiated. What is the nature of this effect on positioning at the level of protein–protein interactions? An important insight comes from the observation that the interactions of the signal sequence with elements of the core translocation channel, specifically Sec61α and Sec61β, predict the ultimate timing of cleavage and thus the efficiency of glycosylation. Although one cannot deduce the exact structure of the signal sequence-translocon interaction from cross-linking studies, these studies do reveal a qualitative difference in the nature of this interaction between signal sequences. The cross-linker used in this work is a homobifunctional cysteine-reactive reagent. Notably, SN → QT shows increased cross-linking to both PDI and Sec61β (Fig. 7C).

Fig. 7. Cross-linking to Sec61α and Sec61β predicts maturation efficiency. A, Prl and SN → QT truncated at the indicated chain lengths were synthesized in the presence of RM, and targeted chains were isolated by sedimentation. Resuspended material was set aside or treated with 1 mM 1,8-bis-maleimidotriethylenglycol for 60 min, followed by quenching of the cross-linker and SDS-PAGE. The migration of uncross-linked material (un-XL; i.e. material that did not react with the cross-linker) is indicated, as are the positions of several species that consistently vary between the two substrates. B, Prl and SN → QT were targeted at a chain length of 70 amino acids and subjected to cross-linking as in A. SN → QT-favoring cross-links of 10 and 50 kDa are indicated, in addition to p35 and p45. Uncross-linked material is shown as a separate panel because it was taken from a slightly lighter exposure to demonstrate the equivalence of loading. C, Prl and SN → QT were subjected to cross-linking as in B, and the cross-links were immunoprecipitated with antibodies against PDI, Sec61α, Sec61β, or a non-immune serum. D, Prl, SN → NH, SN → MP, and SN → QT truncated at 70 amino acids were crossed with Sec61α or cross-linker, and the cross-links were immunoprecipitated with antibodies against Sec61β as in C. The ratio of cross-links to Sec61β alone (i.e. p10) and Sec61α-Sec61β together (i.e. p45) were quantitated and expressed as a fraction. E, the Prl and SN → QT signal sequences were fused to the mature domains of TIGRp and GH. The substrates were truncated at a chain length of 70 amino acids and subjected to cross-linking as in B. Uncross-linked material and cross-links to Sec61α alone and Sec61α-Sec61β together are indicated. For both TIGRp and GH substrates, the single cysteine residue is located in the signal sequence; none is present in either mature domain at the chain length used.

Whereas the glycosylation site which we engineered into the Prl substrate was placed at a position where relative proximity to the ER membrane might be an issue, the interaction between signal sequence cleavage and glycosylation is probably more complex, at least in this case. The signal sequences used in this study have similar effects on the glycosylation of both TIGRp and PrlCHO (Fig. 3, B and C). The glycosylation acceptor site in TIGRp, however, is located 25 amino acids away from...
the site of signal sequence cleavage, making it unlikely that simple proximity of the site to the ER membrane prevents its recognition by the glycosylation machinery. In addition, whereas signal sequence cleavage is required for efficient glycosylation of Prl CHO, a small but significant amount of glycosylation (~20–30%) can clearly occur independent of signal sequence cleavage (Fig. 5D).

It has already been documented that topogenic sequences in a protein can act at a distance to modulate recognition of the nascent chain by the glycosylation machinery (34). We propose that, prior to its cleavage, the signal sequence leaves the nascent chain in a conformation that is refractory to efficient glycosylation. Cleavage allows glycosylation to occur, not by simple liberation of the N terminus from the plane of the ER membrane, but by altering the conformation of the N terminus in such a way that it can be recognized and modified. If, however, this cleavage event takes place after a substantial amount of chain elongation into the lumen has occurred, then even signal sequence cleavage is insufficient to allow the chain to assume a favorable conformation for glycosylation. Chains that undergo delayed cleavage (or no cleavage at all) are capable only of the minor amount of glycosylation that presumably occurs at some later point during or after translocation. It will be interesting to see whether the recently reported effect of a signal sequence polymorphism on CTLA-4 glycosylation (35) is via a similar mechanism.

Our work adds an additional role to the growing list of posttranslational functions for signal sequences. In addition to stimulating the initiation of translocation, a functional requirement of all signal sequences, signal sequences also influence the dependence of a protein on the translocating chain-associated membrane protein (27) or the translocon-associated protein complex (13) for successful translocation. Signal sequences regulate the nature of the interaction between the ribosome and translocon and the timing with which a translocating nascent chain has access to the ER lumen (14–16). From this study, we conclude that the signal sequence also is responsible for setting in motion the pathway of the nascent chain through maturation in the ER lumen. In this case and others (36), the influence of a signal sequence is exerted through the timing of its cleavage, but we see no reason why this need always be the case. It is tempting to speculate that the incredible diversity seen among signal sequences (37) is a consequence of evolution having optimized signal sequences to faithfully guide the biogenesis, modification, and folding pathways of secretory and transmembrane proteins (38). Indeed, it may become apparent that the role of signal sequences in guiding protein folding at the ER is every bit as critical as their role in mediating targeting and the initiation of translocation.

Acknowledgments—We thank A. Calayag, C. Zimmerman, M. Corpuz, N. Singh, and S. Cachaco for technical assistance; R. S. Hegde for stimulating discussions; and J. E. A. Michaels, S. Saghafi, and X. Shen for critical comments on the manuscript.

REFERENCES

1. Keenan, R. J., Freyman, D. M., Stroud, R. M., and Walter, P. (2001) Annu. Rev. Biochem. 70, 755–775

2. Walter, P., and Johnson, A. E. (1994) Annu. Rev. Cell Biol. 10, 87–119

3. Görlich, D., Hartmann, E., Prehn, S., and Rapoport, T. A. (1992) Nature 357, 47–52

4. Görlich, D., Prehn, S., Hartmann, E., Kalies, K. U., and Rapoport, T. A. (1992) Cell 71, 489–503

5. Enssault, Y., Feldheim, D., Blondel, M. O., Schekman, R., and Kepes, F. (1994) J. Biol. Chem. 269, 27478–27485

6. Kalies, K. U., Rapoport, T. A., and Hartmann, E. (1998) J. Cell Biol. 141, 887–894

7. Levy, R., Wiedmann, M., and Kreibich, G. (2001) J. Biol. Chem. 276, 23450–23456

8. Jungnickel, B., and Rapoport, T. A. (1995) Cell 82, 261–270

9. Mothes, W., Jungnickel, B., Brunner, J., and Rapoport, T. A. (1998) J. Cell Biol. 142, 355–364

10. Hegele, R. S., and Lingappa, V. R. (1996) Cell 85, 217–228

11. Hamman, B. D., Hendershot, L. M., and Johnson, A. E. (1998) Cell 92, 747–758

12. Crowley, K. S., Lao, S., Sorrell, V. E., Reinhardt, G. D., and Johnson, A. E. (1998) Cell 78, 461–471

13. Fons, R. D., Bogert, B. A., and Hegde, R. S. (2003) J. Cell Biol. 160, 529–539

14. Kim, S. J., and Hegde, R. S. (2002) Mol. Biol. Cell 13, 3775–3786

15. Kim, S. J., Mitra, D., Salerno, J. R., and Hegde, R. S. (2002) Dev. Cell 2, 207–217

16. Rutkowski, D. T., Lingappa, V. R., and Hegde, R. S. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 7823–7828

17. Kim, S. J., Rahbar, R., and Hegde, R. S. (2001) J. Biol. Chem. 276, 26132–26140

18. Johnson, A. E., and van Waes, M. A. (1999) Annu. Rev. Cell Dev. Biol. 15, 799–842

19. Ellgaard, L., and Helenius, A. (2001) Curr. Opin. Cell Biol. 13, 431–437

20. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York

21. Chou, P. Y., and Fasman, G. D. (1974) Biochemistry 13, 222–245

22. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 5, 105–132

23. Stone, E. M., Fingert, J. H., Alward, W. L., Nguyen, T. D., Polansky, J. R., Sunden, S. L., Nishimura, D., Clark, A. F., Nyutuen, A., Nichols, B. E., Mackey, D. A., Ritch, R., Kalenak, J. W., Craven, E. R., and Shefdief, V. C. (1997) Science 275, 668–670

24. Connolly, T., Collins, P., and Gilmore, R. (1989) J. Cell Biol. 108, 299–307

25. Lingappa, V. R., Devillers-Thierry, A., and Blobel, G. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 2432–2436

26. Wolin, S. L., and Walter, P. (1993) J. Cell Biol. 21, 1211–1219

27. Voigt, S., Jungnickel, B., Hartmann, E., and Rapoport, T. A. (1996) J. Cell Biol. 134, 25–35

28. Martoglio, B., and Dobberstein, B. (1998) Trends Cell Biol. 8, 410–415

29. Chen, X., VanValkenburgh, C., Liang, H., Fang, H., and Green, N. (2001) J. Biol. Chem. 276, 2411–2416

30. Gilmore, R., and Blobel, G. (1985) Cell 42, 497–505

31. Geder, V., Bieri, C., and Spiess, M. (1999) J. Cell Biol. 147, 257–266

32. Nilsson, I., and von Heijne, G. (1993) J. Cell Biol. 100, 5798–5801

33. Pastewl, M., Dubey, R. E., and Strynadka, N. C. J. (1998) Nature 396, 186–190

34. Nilsson, I., and von Heijne, G. (2000) J. Biol. Chem. 275, 17338–17343

35. Anjose, S., Nguyen, A., Oumissi-Benkalha, H., Tessier, M., and Polychronakos, C. (2002) J. Biol. Chem. 277, 44478–44486

36. Li, Y., Luo, L., Thomas, D. Y., and Kang, C. Y. (2000) Virology 272, 417–428

37. von Heijne, G. (1985) J. Mol. Biol. 184, 99–105

38. Lingappa, V. R., Rutkowski, D. T., Hegde, R. S., and Andersen, O. S. (2002) Bioessays 24, 741–748
Signal Sequences Initiate the Pathway of Maturation in the Endoplasmic Reticulum Lumen

D. Thomas Rutkowski, Carolyn M. Ott, Jon R. Polansky and Vishwanath R. Lingappa

J. Biol. Chem. 2003, 278:30365-30372.
doi: 10.1074/jbc.M302117200 originally published online May 27, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M302117200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 35 references, 17 of which can be accessed free at
http://www.jbc.org/content/278/32/30365.full.html#ref-list-1