

1998-09-18

Transmembrane protein insertion orientation in yeast depends on the charge difference across transmembrane segments, their total hydrophobicity, and its distribution

Carol A. Harley
Albert Einstein College of Medicine

Jonathan A. Holt
University of Massachusetts Medical School

Rhiannon Turner
Cambridge University

See next page for additional authors

Follow this and additional works at: https://escholarship.umassmed.edu/maps_pubs

 Part of the [Biochemistry, Biophysics, and Structural Biology Commons](#), [Microbiology Commons](#), and the [Physiology Commons](#)

Repository Citation

Harley, Carol A.; Holt, Jonathan A.; Turner, Rhiannon; and Tipper, Donald J., "Transmembrane protein insertion orientation in yeast depends on the charge difference across transmembrane segments, their total hydrophobicity, and its distribution" (1998). *Microbiology and Physiological Systems Publications and Presentations*. 67.
https://escholarship.umassmed.edu/maps_pubs/67

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Microbiology and Physiological Systems Publications and Presentations by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.

Transmembrane protein insertion orientation in yeast depends on the charge difference across transmembrane segments, their total hydrophobicity, and its distribution

Authors

Carol A. Harley, Jonathan A. Holt, Rhiannon Turner, and Donald J. Tipper

Keywords

transmembrane protein insertion, yeast, endoplasmic reticulum

Rights and Permissions

© 1998 by The American Society for Biochemistry and Molecular Biology, Inc. Publisher PDF posted after 12 months as allowed by the publisher's author rights policy at <http://www.jbc.org/site/misc/edpolicy.xhtml#copyright>.

Transmembrane Protein Insertion Orientation in Yeast Depends on the Charge Difference across Transmembrane Segments, Their Total Hydrophobicity, and Its Distribution*

(Received for publication, April 15, 1998, and in revised form, July 7, 1998)

Carol A. Harley‡, Jonathan A. Holt, Rhiannon Turner§, and Donald J. Tipper¶

From the Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts 01655

The determinants of transmembrane protein insertion orientation at the endoplasmic reticulum have been investigated in *Saccharomyces cerevisiae* using variants of a Type III (naturally exofacial N terminus (N_{exo})) transmembrane fusion protein derived from the N terminus of Ste2p, the α -factor receptor. Small positive and negative charges adjacent to the transmembrane segment had equal and opposite effects on orientation, and this effect was independent of N- or C-terminal location, consistent with a purely electrostatic interaction with response mechanisms. A 3:1 bias toward N_{exo} insertion, observed in the absence of a charge difference, was shown to reflect the N_{exo} bias conferred by longer transmembrane segments. Orientation correlated best with total hydrophobicity rather than length, but it was also strongly affected by the distribution of hydrophobicity within the transmembrane segment. The most hydrophobic terminus was preferentially translocated. Insertion orientation thus depends on integration of responses to at least three parameters: charge difference across a transmembrane segment, its total hydrophobicity, and its hydrophobicity gradient. Relative signal strengths were estimated, and consequences for topology prediction are discussed. Responses to transmembrane sequence may depend on protein-translocon interactions, but responses to charge difference may be mediated by the electrostatic field provided by anionic phospholipids.

The vast majority of eukaryotic integral transmembrane (TM)¹ proteins are co-translationally inserted into the membrane of the endoplasmic reticulum (ER) in a signal recognition particle (SRP)-dependent manner. Insertion occurs by lateral translocation of the TM segments from the translocon complex into the lipid bilayer (1–3). With few known exceptions (4, 5), this results in a unique topology, determined during this insertion process by interaction between topogenic signals within

the nascent protein and response mechanisms that are, at best, only partially understood.

TM proteins with single TM segments may be classified according to the initial location of their N terminus and the presence or absence of a cleaved N-terminal signal sequence. Type I TM proteins have a cleaved N-terminal signal sequence followed by a TM anchor (stop transfer) segment, and so they are initially inserted N_{cyt} , although the mature N terminus is N_{exo} . Type II proteins contain an uncleaved signal/anchor sequence resulting in an N_{cyt} orientation, whereas Type III proteins are unique in being inserted N_{exo} , so that the native N terminus is translocated to the lumen. TM segments, including uncleaved signal-anchor segments, usually comprise a stretch of 18 or more generally hydrophobic residues, whereas the hydrophobic segment of cleaved signal sequences is distinctly shorter (7–15 residues) and is usually preceded by a short, positively charged N-terminal segment (6).

The signals determining TM protein insertion orientation have been extensively studied in prokaryotes, primarily by mutational analysis of leader peptidase (7, 8). The primary signal is provided by positively charged residues adjacent to the TM segment; the terminus with adjacent positive charges is normally retained in the cytoplasm, a pattern codified in the “positive inside rule” (9). A major determinant of this pattern is the prokaryotic negative-inside transmembrane potential (10). The presence of several positive charges on both sides of a TM segment can prevent insertion, causing “topological frustration” (11), although it has recently been shown that reducing the anionic phospholipid content of the *Escherichia coli* cytoplasmic membrane facilitates translocation of positive charges in *E. coli* (12). Although the positive inside rule discounts the importance of negatively charged residues (8), negative charges in the short N terminus of the 44-residue phage pf3 coat protein do promote its translocation (13).

Proteins are translocated by several mechanisms in both prokaryotes and eukaryotes. Some prokaryotic proteins require the Sec pathway defined by *sec* genes, others do not, and TM segments of the same protein can belong to both groups (7). Exofacial N termini, as in the pf3 coat protein, belong to the Sec-independent group (14). If the topogenic effects of both transmembrane potential and lipid composition result from purely electrostatic interaction with charges in translocated proteins, these effects may be independent of the translocation mechanism. A minority of proteins in yeast are preferentially translocated at the ER by the SRP-independent post-translational pathway. The known substrates are secreted proteins with signal sequences of unusually low hydrophobicity or Type I TM proteins with weakly hydrophobic N-terminal signal segments, such as Kex2p (15). All other TM proteins, including all of the variants of the model proteins described in this paper, are predicted to be preferentially translocated by the SRP-de-

* This work was supported by Grant 3777A (to D. J. T.) by the Council for Tobacco Research. 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.

‡ Present address: Dept. of Developmental and Molecular Biology, Albert Einstein College of Medicine, Bronx, NY 10461.

§ Present address: Dept. of Biochemistry, Cambridge University, Cambridge CB2 1QP, United Kingdom.

¶ To whom correspondence should be addressed. Tel.: 508-856-2308; Fax: 508-856-5920; E-mail donald.tipper@ummed.edu.

¹ The abbreviations used are: TM, transmembrane; bp, base pair(s); ER, endoplasmic reticulum; N_{cyt} , cytoplasmic N terminus; N_{exo} , exofacial N terminus; PCR, polymerase chain reaction; SRP, signal recognition particle; HN, hemagglutinin-neuraminidase; KD, Kyte-Doolittle; GES, Goldman, Engelman, and Steitz.

pendent pathway (15). As for prokaryotes, mechanisms for response to charged topogenic signals in eukaryotes may be independent of the pathway.

Although the positive inside pattern also tends to be true for eukaryotic TM proteins of known topology, statistical analysis showed a better correlation with the charge difference across a TM segment, calculated for an arbitrary window of 15 residues on either side (16), and a much greater tolerance for translocation of positive charges is apparent. Positive values for $\Delta(C - N)$, the difference between C-terminal and N-terminal charges, correlate with N_{exo} orientation, and negative values correlate with N_{cyt} orientation. A bias toward N_{exo} insertion is seen when $\Delta(C - N)$ is zero (16). This "charge difference rule" differs from the positive inside rule by giving positive and negative charges equal weight. There is no detectable pH gradient at the ER to respond to this signal. The implied difference in response, therefore, may reflect differences in the mechanisms determining orientation. Results of a limited *in vitro* analysis of the effects of altering flanking charge on the topology of cytochrome P450, a Type III protein with a very short N-terminal segment, were consistent with this charge difference rule (17), whereas *in vivo* analyses of the asialoglycoprotein receptor subunit H1 and paramyxovirus hemagglutinin-neuraminidase (HN), both Type II TM proteins, were interpreted as indicating compliance with the positive inside rule, with a dominant effect for N-terminal charge (18–20). More recently, however, it has become apparent that Type II proteins with N termini of significant length may be poor models for tests of topogenic signals because, even in cotranslational translocation, there are no necessary constraints on folding in this N-terminal segment prior to engagement of the first TM segment with the SRP complex. Stable folding can be a strong impediment to reversal of orientation, providing an N_{cyt} bias, as shown for variants of the asialoglycoprotein receptor (21). In contrast, cytoplasmic folding stability in the N-terminal segment of a Type III protein must be constrained to allow its subsequent translocation; use of such a protein in analysis of topogenic signals should avoid the bias inherent in many Type II proteins.

G protein-coupled receptors have an exofacial (N_{exo}) N terminus and seven TM segments, so that ER insertion of their first TM segment is topologically equivalent to insertion of a Type III protein. This is true however, only if the N-terminal segment is less than about 100 residues. Above this size limit, an N-terminal signal is employed and insertion is analogous to that of a Type I protein (22). We have constructed model fusion proteins (see Fig. 1) based on N-terminal fragments of Ste2p, the yeast G protein-coupled α -factor pheromone receptor, which has typical topology and an N_{exo} 51-residue N terminus (23). The effects of alterations in flanking charge on insertion orientation of the first TM segment (TM_1) of Ste2p in these fusions was found to be entirely consistent with the charge difference rule (24). Although these results indicated that N-terminal negative charges provide strong topogenic signals, the relative strength of positive and negative charges was not determined, and their effectiveness in N- and C-terminal location, relative to the TM segment, was not compared. We have now compared the topogenic effects of small N- and C-terminal positive and negative charges, the first such analysis in any model system. An improved assay procedure increases the precision of our *in vivo* orientation analysis. Precise equivalence of charges, independent of polarity and location, *i.e.* strict compliance with the charge difference rule, was demonstrated.

Earlier data also indicated that our model fusion proteins had a charge-independent bias toward N_{exo} insertion of about 3:1 in the absence of any charge difference (24). Current data

confirm this bias and indicate that neither the 51-residue N-terminal segment nor the C-terminal reporter affect this charge-independent bias, the cause of which, therefore, probably lies in the TM segment itself. *In vitro* studies of a TM protein with a very short N-terminal domain and an oligoleucine TM segment showed that a longer TM segment favors N_{exo} insertion, whereas a short segment favors N_{cyt} insertion (25). In a mutant H1 receptor subunit in which the N-terminal domain was reduced to 4 residues, but in which the -3 charge difference and the normal N_{cyt} orientation were retained (18), orientation was normal when the TM segment was replaced by 7–16 leucines but was partially reversed when this was increased to 19 or more leucines (26). With a charge difference of -2 , topology remained predominantly N_{cyt} with 7–13 leucines but was almost completely reversed in constructs with 19–25 leucines (26). Topogenic effects of TM segment length, therefore, can override a small charge difference signal. No studies have been reported in the absence of a charged topogenic signal or that distinguish effects of length from those of total hydrophobicity. We have now shown that TM segment hydrophobicity (rather than length) is largely responsible for the observed charge-independent bias in orientation of our model Type III proteins and have deduced an estimate of relative signal strengths. We also find that the distribution (gradient) of hydrophobicity within the TM segment is a previously unidentified and significant determinant of orientation that could not have been detected in analyses of homopolymeric TM segments (25, 26). The consequences for topology prediction are discussed.

MATERIALS AND METHODS

Strains, Reagents, and Assays—All DNA manipulations were performed using the *E. coli* strain DH5 α (*supE44D lacU169*(β 80 *lacZDM15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*). Most fusion proteins were expressed in *Saccharomyces cerevisiae* strain CRY2A, a derivative of strain CRY2 (*MAT α ura3-1 leu2-3,112 trp1-1 his3-11 ade2 can1-100*) (27), in which ER-KEX2 (28), under control of the *PGK* promoter, is integrated at *his3-11*. Many strains were also expressed in the isogenic *MAT α* strain CRY1 (27) grown under Leu selection to ensure retention of the episomal YEp(*LEU2*)ER-KEX2 plasmid (28), in which ER-KEX2 is expressed from the *KEX2* promoter. Cell fractionation, β -lactamase activity assays, and Western blots using anti- β -lactamase sera or antisera to the Ste2p N terminus were performed as described previously (24). Insertion orientation, expressed as % N_{cyt} , was deduced from β -lactamase activity distribution in strain CRY2A, where A_{sec} = the fraction of total activity secreted and A_{ca} = the fraction cell-associated ($A_{\text{sec}} + A_{\text{ca}} = 1$), using the formula % $N_{\text{cyt}} = 100(A_{\text{sec}}/0.85)/2(A_{\text{ca}} - 0.15A_{\text{sec}})$. This is based on half-lives of 4 and 8 h for A_{ca} and A_{sec} activities, respectively, and the observation that 15% of secreted activity is cell-associated by virtue of being bound to the cell wall or in transit (see under "Results"). This is a revised form of the formula previously used (24) incorporating significantly smaller correction factors. Antibody to Ost1p was kindly provided by Dr. Reid Gilmore.

DNA Manipulations - Construction of Charge Mutants—Constructs S79g-PB and S42-l, m, n, o, p, q, r, and s have been previously described (24) and are renamed as constructs 5, 6, 10, 20, 4, 9, 23, 24, and 27, respectively. The Ste2p fragments in all S42 and S79 constructs are bounded by N-terminal *XhoI* and C-terminal *PstI* sites and contain an *NdeI* site near the center of TM_1 . The interchangeable P or α fragments of the reporter are bounded by *PstI* and *MscI* sites. (24). All of these sites and a vector *AatII* site are unique and were used to interchange these fragments in many of the constructs described here. The sequences of all products were confirmed.

All new constructs were derived from S42h-PB (24), used as template in a PCR with the N-terminal primer 5'-GGGCTCGAGAATG-AACRACCAGTTGCAAGGTTTGTAGTTAAAC (where R = A or G) and C-terminal primer 5'-GAGTATGGCTGCAGTCGG, producing a mixed pool of 120-bp products. These were cut with *XhoI* + *PstI* (sites underlined) and cloned into a YEp-*LEU2* derivative of S79a-PB (24) also cut with *XhoI* + *PstI*, producing S42-PB construct 1. In all S42 fusions, the Ste2p TM_1 segment is preceded by a 15-residue N-terminal sequence that varies only in residues 2–4; in construct 1, residues 1–4 are MNNQ. In all fusions, the segment 3–6 residues downstream of TM_1 comprises the variable C-terminal component. In construct 1, this se-

N-terminal Flanking sequences from Ste2p

S79 **M**SDAAPSLNSLNFYDPTYNPGQSTI**N**YTSIYG**M**GSTIT**F**DELOGLVNSTV**T**Q
 S72 **M**DPYINPGQSTI**N**YTSIYG**M**GSTIT**S**S**M**NOLOGLVNSTV**T**Q
 S60 **M**DPYINPGQSTI**N**YQSS**M**NOLOGLVNSTV**T**Q
 S42 **M**DD**E**LOGLVNSTV**T**Q

Ste2p TM₁ transmembrane segment (R58I) AIMFGVICGAAALTLIVM**M**

The Ste2p 8 residue C-terminal flanking sequence and PB or αB reporter

S79-PB **T**SR**S**R**K**R**K**T**P**•**A**AIL**P**A**N**ASQ**V**Y**K**R*****SD**T**A**E**H**T**T**S**F**V**V**T**N**N**Y**T**SL**H**T-
 DL**I**H**H**G**N**G**T**Y**T**T**F**T**T**PH**I**PA**V**A**K**R*****DL**R**•**β**la
 S60-αB **S**ST**S**T**S**DT**T**•**A**AV**S****N**ST**N**GL**L**F**I**NT**I**AS**I**A**A**KE**G**V**A**K**R**•**DL**R•**β**la

FIG. 1. **Model Type III TM fusion proteins.** Charged residues in these Ste2p-β-lactamase fusions are shown in **boldface**, and Asn residues that are known to be N-glycosylated, when luminal, are *outlined*. Unless otherwise indicated, all fusions contain TM₁, the 20-residue Ste2p first transmembrane segment in which the seventh residue, Arg in native Ste2p, has been mutated to Ile (R58I). Each protein is made up of TM₁ flanked by one of the indicated N-terminal sequences and variants of the 8-residue Ste2p first cytoplasmic loop. This is followed by the PB or αB reporter in which B is β-lactamase (*bla*), the mature 29-kDa β-lactamase sequence. ● indicates the junctions between Ste2p, P, or α- and β-lactamase fusion components, and the * indicates potential sites of processing by Kex2p. The 15 flanking residues that contribute to the calculated Δ(C - N) signal are *underlined*. N-terminal charges are varied in residues 2–4 of S42 fusions. C-terminal charges are varied only in the **RSRK** sequence within the Ste2p loop preceding PB or αB by conversion of the positively charged residues to Thr or Asp. The sequences shown preceding PB and αB illustrate the widest variation in C-terminal charge among analyzed fusions.

quence is RTRK. Construct 1 was cut with *NdeI* + *AatII* and the 4146-bp fragment containing *LEU2* and the Ste2p N-terminal sequence was ligated to a 2905-bp *NdeI* + *AatII* fragment derived from S79b-αB (24). The product, S42-αB construct 21, has the N-terminal MNNQ sequence and the C-terminal TTTT sequence. The 4146-bp *NdeI* + *AatII* fragment from construct 21 containing the Ste2p N-terminal MNNQ sequence and the *LEU2* marker was ligated with the 2905-bp *NdeI* + *AatII* fragment from S79f-PB (24) encoding the C-terminal sequence TTRT producing S42-αB construct 11. Construct 21 was used as a template, with the N-terminal degenerate primer 5'-GGGCTC-GAGAATGAACRACCAGTTGCAAGGTTAGTTAAC (where R = A or G) and the C-terminal primer 5'-GGGCTGCAGTCGGGTCTGTTGTGCTTGTGCGATGTC to generate a mixed pool of ~120-bp PCR products, which were cut with *XhoI* + *PstI* and recloned into the corresponding sites of construct 21. Sequence analysis identified construct 28 in which the C-terminal amino acid sequence has been mutated to TTTD.

Primers 5'-GGGCTGCAGTCGGGTCTGTTGTGCTTGTGCGATGTC and 5'-GGGCTCGAGAATGAACGACCAGTTGCAAGGTTAGTTAAC were used in a PCR with construct 28 as a template to produce a 120-bp product. This was cut with *XhoI* + *PstI* and cloned into construct 28. The resultant constructs 25 and 26 have the MNDD N-terminal and TTTD C-terminal sequences. Constructs 11 and 25 were cut with *NdeI* + *AatII*. The 4146-bp fragment from 25 encoding the N-terminal sequence MNDQ was ligated to the 2905-bp fragment from 11 encoding a C-terminal TTRT fragment, resulting in constructs 7 and 8. Constructs 13 and 14, 2, 3, and 15 were made using a similar strategy; the N/C-terminal sequences flanking the Ste2p TM₁ segment are MNDQ/TTTT, MNDQ/TTRK, MDNE/TTRT, and MDNE/TTDT, respectively.

S72 and S60 derivatives—The N-terminal primer 5'-GGGCTCGAG-CATGAACCAGTTGCAAGGTTAGTTAACACTAGTGTACTCAGGC-CATTATG and the C-terminal primer 5'-GGGCTGCAGTCGGTGTG-TGCTTGTGCGACGACATCCACATGACATTC were used with construct 21 as template to produce a 120-bp product, which was cloned as a *XhoI* to *PstI* fragment into construct 21 producing construct 22. The sequence at the *XhoI* site is modified such that the neutral SSS tripeptide is in frame with the downstream initiator M residue. S79g-PB was used as a template for PCR using primers 5'-GGGCTCGAGAATGRATCCAACG-TATAATCCTGG and 5'-GGGCTCGACGAAGTGATGGTAGATCCA-TTC (where R = A or G). The 94-bp PCR products were digested with *XhoI* + *SalI* and cloned into construct 22 digested with *XhoI*. Constructs 16 and 17, which encode S72 derivatives of Ste2p with extended N-terminal sequences (Fig. 1), were identified by sequence analysis. The complementary oligonucleotides 5'-TCGAGAATGRATCCAACGT-ATAATCCTGGTACCAGCACCATTAACTACCAGTCG and 5'-TCGAC-GACTGGTAGTTAATGGTGTGCTGGTACCAGATTATACGTTGGAT-YCATTC (where R = A or G and Y = T or C) were annealed to generate

XhoI and *SalI* sticky ends. This linker was ligated into construct 22 cut with *XhoI*. The resultant constructs, 18 and 19, encode S60 Ste2p derivatives with an extended neutral non-glycosylated N terminus (Fig. 1).

TM₁ Domain Derivatives of S60-PB—All TM₁ domain derivatives are based on the S60-n20 construct 18. The 20-residue TM₁ segment of Ste2p encoded by S60-n20 was initially truncated by cutting construct 18 with *SpeI* + *NdeI* and ligation with linkers produced by annealing oligonucleotide pairs 5'-CTAGTACTCAGAGCA + 5'-TATGCCTGAGTA and 5'-CTAGTACTTTCGGGTGTCA + 5'-TATGACACCGAAAGTA. The resultant constructs, S60-n14 and S60-n17, respectively, were initially identified by the presence of an additional *ScaI* site. S60-n20 was cut with *SpeI* + *SalI* and ligated with the linker formed by annealing oligonucleotides 5'-CTAGTCTGCTAGCTTTGACTTTGATTGTCTAGTGGATGTCG and 5'-TCGACGACATCCACATGACAATCAAAGTCAAA-GCTAGCAGA. The product, S60-n11, was initially identified by its unique *NheI* site (underlined). Leucine residues were added to the N terminus of the TM₁ segment in S60-n11 by cutting with *SpeI* + *NheI* and ligating with linkers produced by annealing oligonucleotide pairs 5'-CTAGTACTCTGTGCTACTG + 5'-CTAGCAGTAGCAACAGAGTA and 5'-CTAGTACTCTCTACTGTTGCTACTG + 5'-CTAGCAGTAGCAACAGTAGGAGAGTA. The resultant plasmids are S60-L₅-n11 and S60-L₅-n11, respectively. To move the L₅ peptide from the N terminus to the C terminus of S60-n11, S60-n11 was used as template in a PCR with the primers 5'-TCGAGAATGGATCCAACG and 5'-GGGCTCGA-CAGGAGTAGGAGCAGCATCCACATGACAATCAAAG. The resultant product was digested with *BamHI* + *SalI* and ligated into the corresponding sites of S60-n20, producing S60-n11-L₅. The oligonucleotides 5'-GGGACTAGTGTTACTCAGATCGCAATCATCTTACTTTTA and 5'-GGGCTCGACGTCATACTATATGCCAGGGC were used in a PCR to produce a 90-bp product encoding the TM₁ segment from NDV HN (29). This PCR product was cut with *SpeI* + *SalI* and ligated into the corresponding sites of S60-n20, producing S60-HN. Reversal of the HN TM segment was achieved by ligating the same S60-n20 fragment with the *SpeI* to *SalI* linker produced by annealing and ligating two oligonucleotides pairs, 5'-CTAGTGTACTCAGATGAGCTACGCCCTGGC-CGCCGCT + 5'-GGCGATGGAGGCGGCGCCAGGGCGTAGCTCA-TCTGAGTAACA and 5'-CCATCGCCTTGACCATGGTTACTCTGCTG-CTGATCGCCATCACG + 5'-CCATCGCCTTGACCATGGTTACTCTG-CTGCTGATCGCCATCACG. The product is S60-HNrev.

RESULTS

An Improved in Vivo Assay for Protein Insertion Orientation—S79-PB is a chimeric model Type III TM protein (Fig. 1), the derivatives of which, expressed from the *PGK* promoter on multicopy YEplasmids, have previously been used to analyze topogenic signals determining insertion orientation in *S. cerevisiae* (24). S79 is the N-terminal 79 residues of Ste2p, including the 51-residue exofacial N-terminal segment, the 20-residue TM₁, and the 8-residue first cytoplasmic loop (Fig. 1). It is fused at its C terminus to PB or αB, reporters of insertion orientation. P is a 59-residue fragment of the precursor of secreted K1 killer toxin (Fig. 1) (30), including two sites efficiently cleaved by Kex2p, a TM protease, the activity of which is normally limited to the lumen of the trans-Golgi (31). B is the mature form of β-lactamase encoded by the *E. coli bla* gene. β-Lactamase is fully active in yeast in cytoplasmic fusions, in cell-associated TM fusions, or when secreted (30). In αB, P is replaced by α, a 31-residue fragment of prepro-α-factor that also includes a site efficiently cleaved by Kex2p (Fig. 1) (24). In principle, therefore, orientation of Ste2p-β-lactamase fusions can be deduced from the ratio of cell-associated activity, due to intact fusions inserted N_{exo} and secreted activity due to fusions inserted N_{cyt} and processed by Kex2p. Cell-associated materials are also routinely monitored for membrane insertion and N-glycosylation by Western blot of membrane fractions, using anti-β-lactamase antibody. If fusions have the predicted size, behave as integral membrane proteins, and have the predicted patterns of N-glycosylation (Fig. 1), this provides independent evidence of insertion orientation (24).

Processing of N_{cyt} fusions by Kex2p normally requires their transport from the ER to the trans-Golgi. Because exit of abnormal TM proteins from the ER may be delayed, we previ-

ously expressed all fusions in cells of yeast *MAT α* strain CRY1 in which additional Kex2p is supplied in the form of ER-Kex2p. ER-Kex2p is a soluble, secreted form of the protease in which the C-terminal TM domain is replaced by an ER retention signal (HDEL). When expressed in strain CRY1 from the weak *KEX2* promoter on a multicopy episomal plasmid (28), β -lactamase secretion efficiencies were not affected for any fusion protein (24). However, this level of ER-Kex2p expression failed to restore expression of α -factor in a *kex2-null* mutant of strain CRY2, a *MAT α* strain isogenic to CRY1 (27). We therefore constructed strain CRY2A in which the *ER-KEX2* gene, expressed from the *PGK* promoter (at least 100-fold stronger than the *KEX2* promoter (32)), is integrated at the *his3-1* locus of strain CRY2. Normal α -factor secretion was restored in the isogenic *kex2-null* strain.² Expression of secreted β -lactamase from most fusion constructs was slightly increased in strain CRY2A, whereas secretion from a few was more significantly enhanced, implying inefficient translocation to the trans-Golgi. Strain CRY2A, therefore, has been used routinely in all subsequent orientation assays.

Calculation of orientation from the ratio of secreted to cell-associated β -lactamase activity requires correction for the half-lives of secreted β -lactamase and cell-associated fusion proteins. Half-lives were determined in strain CRY2A using PB fusions expressed from the *GAL1* promoter by following the decay in activity after a shift from galactose (inducing) to glucose (repressing) medium. The half-life of secreted activity was 8 h in cultures buffered at pH 7. Half-lives of cell-associated N_{exo} fusions were significantly increased in strain CRY2A compared with strain CRY1 (24) and were essentially identical (4–5 h) for multiple representatives of all of the different constructs described here. Correction also has to be applied for secreted β -lactamase in transit or trapped in the cell wall. The later accounts for 5–8% of the secreted activity in both CRY1 and CRY2A strains, but activity in transit is reduced from about 15% in CRY1 to about 10% in CRY2A. Correction factors are, therefore, smaller in strain CRY2A (see under “Materials and Methods”), and the precision of *in vivo* orientation analysis is significantly increased relative to earlier data derived from strain CRY1 (24). Total recovered activities varied no more than 2-fold between constructs and were consistent with the expected expression levels (not shown).

The Topogenic Effect of a Single Net Negative Charge Difference Is Precisely Opposite to That of a Single Net Positive Charge Difference—We tested the effects on orientation of one or two positive or negative charges, situated either N- or C-terminal to the TM₁ segment, primarily using variants of S42-PB or S42- α B. In these fusions, the N-terminal segment of S79 is reduced to 15 residues (Fig. 1), and the N-terminal methionine contributes to the net charge difference, calculated assuming absence of N-terminal modification (see under “Discussion”). Orientation data, expressed as percentage of N_{cyt} insertion, are shown in Fig. 2. Both PB and α B fusions were analyzed for five of these fusions. No significant differences were found, strengthening the prior conclusion (24) that these two reporters are equivalent and both neutral with respect to insertion orientation. Sixteen S42 constructs were expressed in strain CRY2A, and data for strain CRY1 are also reported for 11 of these. Differences were usually small, but secreted activities were significantly higher in strain CRY2A, especially for constructs 11 and 14, implying incomplete Kex2 processing of these fusions in the absence of high levels of ER-Kex2p. Previously reported data for constructs 6, 10, and 20 (24), as well as data for three new constructs (1, 21 and 22), also expressed in

strain CRY1, are included for comparison. These data showed the same pattern as those derived from strain CRY2A.

Cell-associated species were analyzed for membrane association by Western blot analysis using anti- β -lactamase antibody. Low speed supernatants from cell breakage were fractionated into 148,000-g supernatant and pellet (membrane) fractions, with and without prior treatment with 0.1 M carbonate buffer, pH 11.5, which solubilizes peripheral but not integral membrane proteins, or with 1% Triton X-100, which solubilizes both. Western blots of cell-associated material from constructs 3, 28, and 2 (Fig. 2), expressed in strain CRY2A, are shown in Fig. 3A. In each case, essentially all of the fusion protein detected was found in the pH 11.5 pellet (Fig. 3A, lanes 3, 7, and 13) rather than in the supernatant (lanes 4, 8, and 14). Because all were also solubilized by 1% Triton X-100 (lanes 2, 6, and 12), all are integral membrane proteins. N_{exo} -oriented S42-PB fusion proteins are predicted to be 40.5-kDa integral membrane proteins, unaffected by treatment with endoglycosidase H, because they have N-glycosylation sites only in P (Fig. 1). This was demonstrated for all (not shown), providing independent confirmation of orientation of this cell-associated material. The identity of these fusions was confirmed by probing with antibody to the Ste2p N terminus, which detected the same species in the pellet fractions, as shown for construct 2 (Fig. 3B, lane 15). The detection of the 66-kDa oligosaccharyltransferase (Ost1p) in a representative pellet fraction (Fig. 3C, lane 17) served to demonstrate that this fraction included the ER membrane. These blots were representative of those obtained with all of the fusions; in each case, the cell-associated fusion protein was present exclusively as an N_{exo} integral membrane protein.

In pellet fractions from cells expressing S42-PB constructs with predominantly N_{cyt} insertion, uncleaved N-glycosylated fusions can only be detected on prolonged exposure, (not shown), consistent with efficient Kex2p cleavage in strain CRY2A. β -Lactamase is visible as a soluble cell-associated 29-kDa species in the supernatant fractions from these same cells (Fig. 3A, lanes 6 and 8) and presumably represents Kex2p-cleaved enzyme in transit. The band is relatively weak, consistent with the short half-life of about 20 min for secretion of β -lactamase (30). The Ste2p fragment released by Kex2p action is 8.7 kDa, detectable using antibody to the Ste2p N terminus, as previously reported (not shown; Ref. 24).

In the four S42 constructs with zero charge difference (Fig. 2, constructs 10–15), this neutrality depends on balanced N- and C-terminal charges or on balance between charges within the N-terminal segment only. Because the effects of such charges may vary with distance from the TM segment, some fractional charge may remain. To eliminate such effects, we constructed S72 fusions in which all of the N-terminal charges, apart from the terminal methionine, were eliminated (Fig. 1). In these neutral S72- α B fusions, the minimal separation of charge from the TM segment is 31 residues. They had the same 25–30% N_{cyt} insertion ratio as the neutral S42 constructs, independent of the presence (construct 17) or absence (construct 16) of an Asp residue adjacent to the N-terminal methionine (Fig. 2). N-terminal N-glycosylation sites in S72 were eliminated to allow C-terminal glycosylation patterns to be used in orientation analysis, producing S60-PB constructs (constructs 18 and 19). As for S72 fusions, orientation was unaffected by a subterminal Asp residue (not shown). Clearly, charges separated from the TM segment by 31 neutral residues do not influence topology. In general, orientation (% N_{cyt}) was consistent within each group defined by a common charge difference (Fig. 2). Thus, a $\Delta(C - N)$ of +2 had the same effect on orientation, independent of whether this resulted from two N-terminal negative charges,

² C. A. Harley and D. J. Tipper, unpublished observations.

Construct #	Type	Flanking sequence Charges			%N _{cyt}	
		N-terminal		C-terminal	CRY1	CRY2
Δ(C-N) = +2						
1.	42-PB	MN··	+1	R·RK	+3	4
2.	42-PB	MND·	0	·RK	+2	2 3
3.	42-PB	MD·E	-1	·R·	+1	3 7
4.	42-PB	MDDE	-2	····	0	5 9
5.	79-PB	MD··DE	-2	····	0	3 4
Average %N_{cyt} 5						
Δ(C-N) = +1						
6.	42-PB	MKKE	+2	R·RK	+3	10
7.	42-αB	MND·	0	·R·	+1	8 13
8.	42-PB	MND·	0	·R·	+1	14
9.	42-PB	MD·E	-1	····	0	8 14
Average %N_{cyt} 13						
Δ(C-N) = 0						
10.	42-PB	MK·K	+3	R·RK	+3	25
11.	42-αB	MN··	+1	·R·	+1	17 31
12.	42-PB	MN··	+1	·R·	+1	44
13.	42-αB	MND·	0	····	0	23 37
14.	42-PB	MND·	0	····	0	19 42
15.	42-PB	MD·E	-1	··D	-1	25
16.	72-αB	MN-····	0	····	0	26
17.	72-αB	MD-····	0	····	0	25 31
18.	60-αB	MD-····	0	····	0	24
19.	60-PB	MD-····	0	····	0	21 22
Average %N_{cyt} 30						
Δ(C-N) = -1						
20.	42-PB	MKKK	+4	R·RK	+3	51
21.	42-αB	MN··	+1	····	0	46
22.	42-PB	MN··	+1	····	0	54
23.	42-PB	MD·K	+1	····	0	73 59
24.	42-PB	MK·E	+1	····	0	58 67
25.	42-αB	MND·	0	··D	-1	42 60
26.	42-PB	MND·	0	··D	-1	46
Average %N_{cyt} 55						
Δ(C-N) = -2						
27.	42-PB	MKKE	+2	····	0	90 85
28.	42-PB	MN··	+1	··D	-1	78
Average %N_{cyt} 81						

FIG. 2. Protein insertion orientation data for fusions with small or zero charge differences. Construct type indicates the length of the fragment derived from Ste2p (S79, S72, S60, or S42) and the reporter to which it is fused (Fig. 1). The N-terminal flanking sequences shown for S42 constructs include the N-terminal methionine and the next three residues; the other 11 are uncharged and remain invariant (Fig. 1). For the S79, S72, and S60 constructs, the equivalently located residues at -15 to -12 are shown (Fig. 1). The S72 and S60 construct N-terminal segments are uncharged except for the N-terminal residues shown. In each case, the C-terminal flanking sequences shown are those from +3 to +6; the rest are uncharged and remain invariant (Fig. 1). All data shown are averages from at least two independent transformants, each assayed at least three times. The variation between experiments was less than 20% of the reported averages. Within each group, defined by a particular net charge difference, the average %N_{cyt} is recorded for the CRY2A data plus the CRY1 data for the six constructs assayed only in that strain.

two C-terminal positive charges, or one of each. Similarly, a Δ(C - N) of +1 had the same effect on orientation, independent of whether this resulted from a single N-terminal negative charge or a single C-terminal positive charge. A similar conclusion can be drawn for the seven constructs with Δ(C - N) of -1, including those with single N-terminal positive or C-terminal negative charges, and for the two constructs with Δ(C - N) of -2. More variation was seen when Δ(C - N) was zero, even among the S42 fusions. Modest variation between the S42, S72, and S60 constructs, especially a small but

consistent reduction in N_{cyt} insertion of S60 constructs, may be attributable to structural differences in the N-terminal segment and was not further investigated. Overall, however, the data indicate a strong correlation of orientation with Δ(C - N), measured over the arbitrary window of 15 residues either side of the transmembrane segment, with equal contributions by either N- or C-terminal charges and equivalent and opposite effects for positive and negative charges. Effects of a small net charge difference were the same whether these resulted from the difference between two substantial charges, as in constructs 6, 10, and 20, or from single charges. These results are consistent with a simple electrostatic interaction between the charged orientation signal and its receptor. The plot of %N_{cyt} against Δ(C - N) indicates a 50/50 distribution when Δ(C - N) is about -1 (Fig. 4). For this particular set of model Type III transmembrane proteins, therefore, the bias toward N_{exo} insertion has a strength close to that of a net Δ(C - N) of +1. The curve is sigmoidal and essentially symmetric about the -1 point, indicating a similar charge-independent bias at all values of the charge difference.

Effect of TM Length and Hydrophobicity on Orientation—Because it has recently been shown that TM segment length and/or total hydrophobicity are important determinants of orientation (26), we studied the effects of variation in these parameters on the N_{exo} bias seen in our model Type III TM proteins. We chose to study effects of variation in the TM segment on orientation in a construct with zero charge difference, where the normal 3:1 N_{exo} bias is most obvious (Fig. 4). The S60-PB construct was chosen because it is devoid of charges that should affect orientation. Its N-terminal Met-Asp dipeptide is separated from the TM segment by a 31-residue neutral sequence that, like the Ste2 peptide from which it is derived, is clearly readily translocated, because insertion is about 80% N_{exo}. S60-PB should, therefore, provide a sensitive context, lacking other known topogenic influences, for analysis of the effects of TM segment hydrophobicity and length on orientation. The initial PB fusion is designated S60-n20 to indicate that it has the normal 20-residue Ste2p TM₁. This TM segment was progressively shortened by N-terminal deletions to produce S60-n17, -n14, and -n11 PB fusions, and the TM segment N terminus was then replaced, in the S60-n11 construct, by relatively hydrophobic Leu₃ or Leu₅ peptides, producing the S60-L₃-n11 and S60-L₅-n11 fusions. We hoped, in this way, to distinguish between effects of length and total hydrophobicity. The TM sequences are shown in Table I, which also lists the molecular masses of the S60-PB constructs and TM segment lengths and total hydrophobicities, calculated using the Kyte-Doolittle (KD) and the Goldman, Engelman, and Steitz (GES) scales. The KD and GES scales assign different relative hydrophobicities to certain residues. The most significant differences for the n20 TM sequence and its variants are that Trp, Thr, Met, and Ser are relatively more hydrophobic, and Val and Leu are less hydrophobic in the GES than in the KD scale. Overall, the GES scale is preferred because of its stronger experimental and theoretical basis (33).

Total recovered activities for these six constructs and half-lives for the n20 and n11 constructs fell within the previously observed range (data not shown), so that orientation data should be of comparable accuracy. Analysis of Western blots showed that the pellet fraction of the cell-associated activities from the n20 to n11 series consisted, in each case, of a fusion protein of the predicted 41–42-kDa size (Fig. 5, lanes 1, 7, 13, and 19), unaffected by treatment with endoglycosidase H (not shown) and, therefore, not N-glycosylated. All proteins were solubilized by Triton X-100 (Fig. 5, lanes 4, 10, 16, and 22).

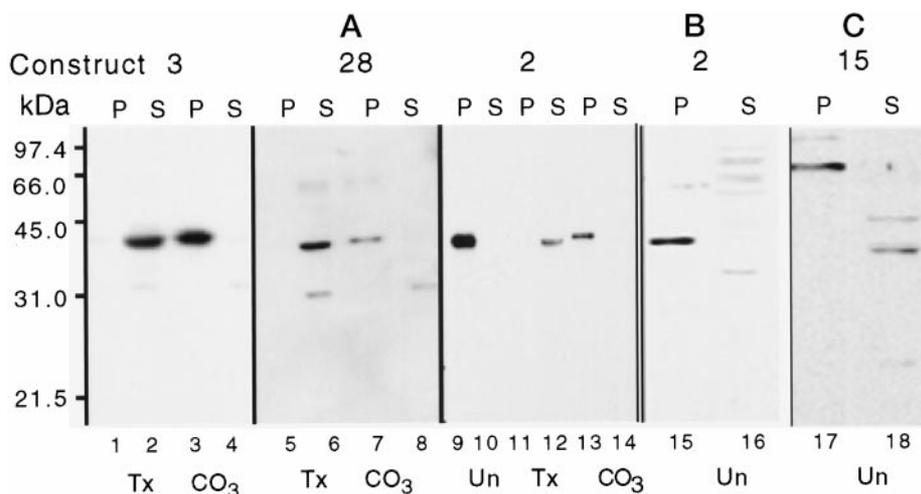


FIG. 3. Western blot analysis of cell-associated materials from cells of strain CRY2 expressing the indicated fusion constructs. All blots represent supernatant (S) or pellet (P) fractions resulting from Airfuge centrifugation of a low speed supernatant fraction from broken cells. This supernatant was either untreated (Un), treated with 1% Triton X-100 (Tx), or treated with 0.1 M Na_2CO_3 , pH 11.5 (CO_3) prior to separation by SDS-polyacrylamide gel electrophoresis and Western blot analysis. A, fractions from cells expressing constructs 3, 28, and 2 (Fig. 1) analyzed with anti- β -lactamase polyclonal antibody. B, untreated fractions from cells expressing construct 2 analyzed with polyclonal antibodies to the Ste2p N terminus. The weakly cross-reactive soluble species were also seen in controls from cells lacking Ste2p fusions (not shown). C, untreated fractions from cells expressing construct 15 analyzed with anti-Ost1p polyclonal antibodies. The soluble 40–45-kDa species are presumably proteolytic fragments of Ost1p.

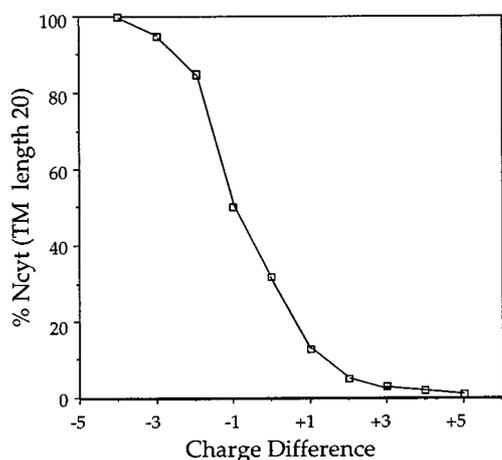


FIG. 4. Relationship of insertion orientation to charge difference. The average $\%N_{\text{cyt}}$ data for the constructs shown in Fig. 2 are plotted against charge difference. Data for charge differences greater than 2 are taken from previously published analyses (24).

Material in the pellet fraction following treatment with pH 11.5 carbonate buffer, therefore, corresponds to an integral membrane protein (Fig. 5, lanes 5, 11, 17, and 23). This fraction comprised essentially all of the cell-associated S60-n20 and S60-n17 fusion proteins (Fig. 5, lanes 5 and 6 and lanes 11 and 12). For the n14 and n11 fusions, however, an increasingly strong signal was seen in the supernatant fraction (Fig. 5, lanes 14 and 20). Because the ratio of fusion protein detected in the pellet and supernatant fractions was not affected by pH 11.5 buffer (Fig. 5, lanes 17 and 18 versus lanes 13 and 14; lanes 23 and 24 versus lanes 19 and 20), the pellet fraction is also an integral membrane protein. Two soluble species were detected, a major 40 kDa species and a minor species of about 58 kDa, also seen in the S60-n17 sample (Fig. 5, lanes 10). The 58-kDa species was converted to a species of 32 kDa on treatment with endoglycosidase H (not shown), and corresponds to the previously observed soluble, glycosylated PB reporter released from N_{cyt} inserted fusions by Kex2p action (24). The 40-kDa species is not glycosylated and therefore corresponds to cytoplasmic, nonintegrated fusion protein. Inefficient insertion presumably

results from severe shortening of the TM segment in the S60-n14 and S60-n11 fusions. This fraction comprised less than 5% of the cell-associated activity except in transformants of the S60-n14 and -n11 fusions, where it was estimated to account for 30 and 65% of the cell-associated activity, or 15 and 25% of total activity, respectively. Although the half-life of this species was not independently tested, other cytoplasmic Ste2p-PB fusions are stable (23), and the half-life of combined cell-associated materials in S60-n11 transformants was normal. $\%N_{\text{cyt}}$ data for these constructs have been adjusted by elimination of this cytoplasmic component (Table I).

Variation of orientation ($\%N_{\text{cyt}}$) with TM segment length follows the previously observed pattern (26): longer, more hydrophobic segments favor N_{exo} insertion, and shorter segments favor N_{cyt} insertion (Fig. 6). When the results from N-terminal insertion of the hydrophobic Leu peptides in the S60-L₃-n11 and S60-L₅-n11 fusions are included, $\%N_{\text{cyt}}$ shows a better correlation with total hydrophobicity, in both the GES and KD scales, than with length (Fig. 6). Thus, the L₅-n11 TM segment, only 16 residues in length, confers an orientation very similar to the n20 TM segment, with an N_{exo} bias distinctly stronger than that of the n17 TM segment.

Effect of TM Segment Hydrophobicity Gradient on Orientation—Hydropathy analysis of the TM segments in Table I indicated marked variation in the distribution of hydrophobicity as well as in its total, as shown for selected constructs in Fig. 7. Distribution is reflected in hydrophobicity gradients, deduced from these plots, and is presented for all of the S60-PB variant TM segments in Table I; positive values favor the N terminus. Because a role for hydrophobicity distribution in determining orientation seemed plausible, additional constructs were made with the intent of testing for such an effect. First, a Leu pentapeptide was inserted at the C terminus of the n11 TM segment, producing the S60-n11-L₅ fusion (Table I). The resultant distinct gradient of hydrophobicity favoring the C terminus reverses the slight gradient in the S60-L₅-n11 fusion (Fig. 7), although length and total hydrophobicity were unchanged. A surprisingly large increase in N_{cyt} insertion resulted (Table I).

The normal 20-residue Ste2p TM₁ segment has a moderate gradient favoring its C terminus (Fig. 7). In order to assess the

TABLE I

Relationship between orientation and TM segment length, total hydrophobicity, and hydrophobicity gradient in derivatives of S60-PB

Derivatives of S60-PB with the indicated TM segments, but otherwise identical, were expressed in strain CRY2A. Total molecular masses are shown. Total hydrophobicities (free energies of transfer from water to oil in kcal/mol) were calculated by summing values for each residue on the KD and GES scales. Gradients were calculated from hydrophilicity plots and linear regression lines, using the GES scale with a window size of 4 residues, as illustrated in Fig. 7. The window size was chosen to minimize the effect of residues adjacent to the TM segment. Because of this effect, however, reversal of the TM segment sequence did not produce a precise reversal of the gradient.

Fusion	Molecule mass	TM segment sequence	Total hydrophobicity		Gradient (GES)	% N _{cyt}
			KD scale	GES scale		
	<i>kDa</i>					
S60-n20	41.9	AIMFG VICGA AALTL IVMWM	45.3	47.5	-3.8	18
S60-n17	41.6	FG VICGA AALTL IVMWM	37.1	39.4	-4.1	32
S60-n14	41.3	ICGA AALTL IVMWM	30.5	32.1	-11	54
S60-n11	41.1	L LALTL IVMWM	27.9	28.4	-6	84
S60-L ₃ -n11	41.4	LLLL LALTL IVMWM	39.3	36.8	0	41
S60-L ₅ -n11	41.7	L LLLLL LALTL IVMWM	46.9	42.4	+1.3	16
S60-n11-L ₅	41.7	L LALTL IVMWM LLLLL	46.9	42.4	-5	56
S60-HN	42.0	IAIL LLTVM TLAIS AAALA	47.2	41.9	+4.2	32
S60-HNrev	42.0	ALAA ASIAL TMVTL LLIAI	47.2	41.9	-7.4	58

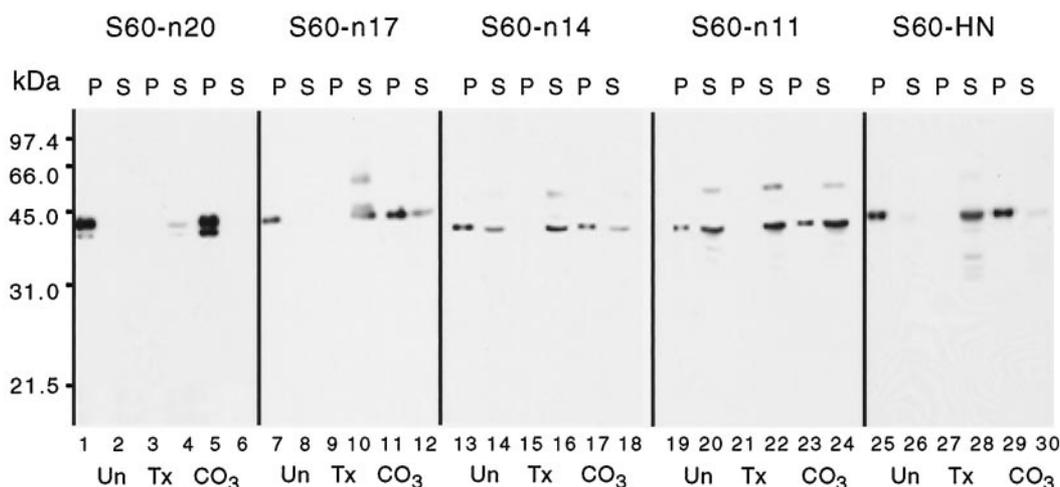
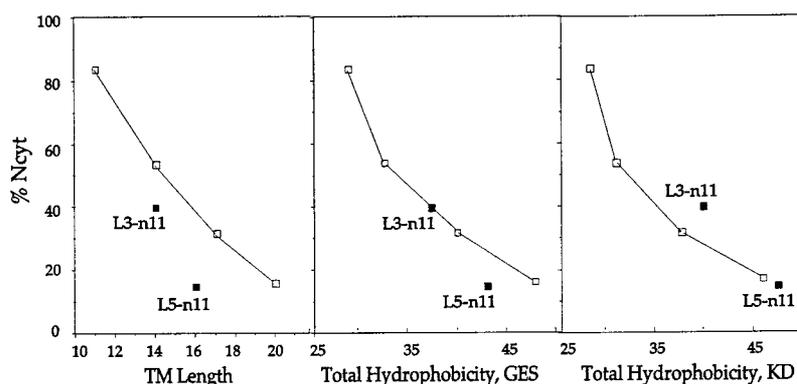


FIG. 5. Western blot analysis of cell-associated material from cells of strain CRY2 expressing variants of S60-PB. All samples were analyzed as in Fig. 3 using anti-β-lactamase polyclonal antibody. Blots represent supernatant (S) or pellet (P) fractions from cells expressing fusion constructs S60-n20, S60-n17, S60-n14, S60-n11, and S60-HN. The proteins observed in the supernatant fractions from fusions with the shortest TM segments (lanes 14 and 20) are cytoplasmic (see under “Results”).

FIG. 6. Relationship of insertion orientation to length or total hydrophobicity of TM segments. Data are plotted for the S60-n20 to S60-n11 series (Table I) (open symbols) and for the S60-L₃-n11 and S60-L₅-n11 constructs (filled symbols).



effect on insertion orientation of an opposite but natural gradient in TM segment hydrophobicity, we replaced TM₁ with the 19-residue TM segment from the Newcastle disease virus HN glycoprotein, producing S60-HN-PB (Table I). The HN glycoprotein is a Type II TM protein with a charge difference of -3, naturally inserted N_{cyt} (29), the TM segment of which has a strong bias in hydrophobicity toward its N terminus by either the KD or GES scale (Fig. 7). For comparison, we also constructed a precise inversion of the HN TM segment producing S60-HNrev-PB (Table I) with the opposite gradient (Fig. 7). All of the cell-associated activity in S60-HN-PB transformants was

present as an integral membrane protein of the predicted 40-kDa size (Fig. 5, lanes 25–30). In keeping with its lower total hydrophobicity (GES scale), N_{cyt} insertion was increased from about 20% for S60-n20 to 32% for S60-HN-PB, the same %N_{cyt} seen for S60-n17, which has the same total hydrophobicity. S60-n17, however, has the opposite gradient (Table I). When the TM segment hydrophobicity gradient was reversed in S60-HNrev-PB, N_{cyt} insertion was increased to 58%. On the GES scale, the entire n20 to n11 series has a gradient favoring the C terminus, so that gradient effects should not invalidate correlations with length or total hydrophobicity. These data show

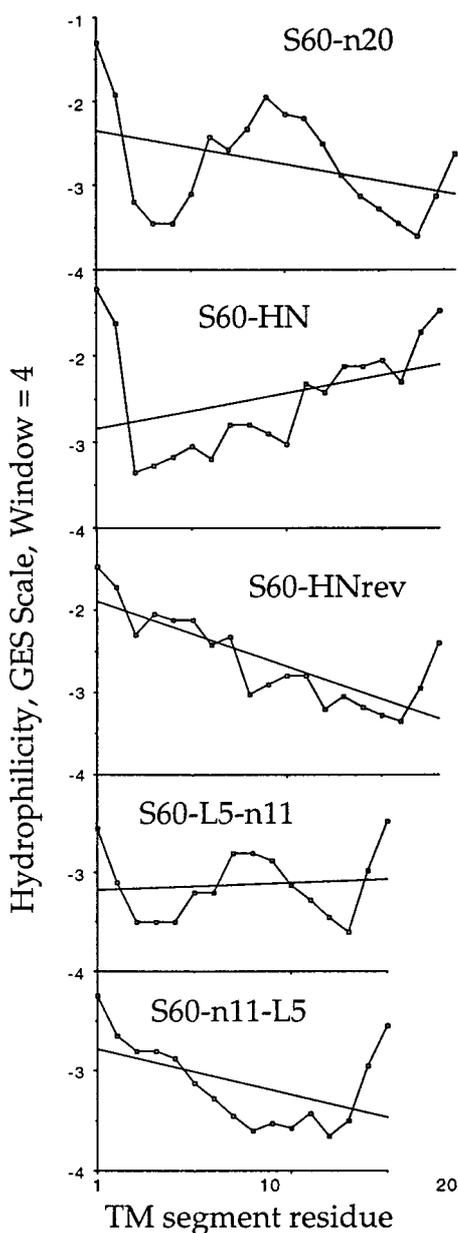


FIG. 7. Hydrophilicity plots for TM segments of five of the S60-PB fusions shown in Table I are plotted using the GES scale with a window of 4 residues. Gradients (Table I) were calculated from the linear regression lines shown.

that TM segment sequence is an important determinant of orientation, independent of total hydrophobicity, and suggest that the more hydrophobic terminus of a TM segment is preferentially translocated.

DISCUSSION

The Charge Difference Orientation Signal—We previously showed (24) that insertion of our model Type III TM proteins at the ER in yeast cells followed the charge difference rule, where charge difference is defined as $\Delta(C - N)$, the arithmetic sum of charged residues within an arbitrary window of 15 residues to either side of a TM segment (16). Using an improved version of this same assay, we have now extended this conclusion by confirming a formal prediction of this rule, that single or double positive and negative charges have equal and opposite effects, independent of N- or C-terminal location relative to the TM segment. $\Delta(C - N)$ is calculated assuming that N-terminal methionines remain unmodified. Although cleavage of methio-

nine would not affect the charge and would not be predicted for any of these constructs (34), co-translational N-acetylation is predicted for cytoplasmic proteins initiating with MD- or MN- (34). Acetylation, however, requires a nascent N terminus of 40–50 residues (34) and, although modification of S72 or S60 fusions may occur (with no predicted effect on topology), it seems unlikely that such modification would occur on the nascent 15-residue N-terminal segment of S42 fusions prior to interaction of TM₁ with the SRP. Although no studies of N-acetylation of TM proteins have been reported, it also seems likely that SRP interaction would mask the S42 N terminus from N-acetylase prior to translocation. Analysis of the status of mature fusions would not resolve this issue, because modification could occur after translocation. The essential *NAT2* gene is required for acetylation at these sites (35), but attempts to use its mutants to study protein synthesis in the absence of Nat2p function have failed.³ Definitive information, therefore, is absent, so that consistency in the topology data is the only available guide. In particular, S42 fusions with the same apparent charge difference as S60, S72 and S79 fusions had similar topologies, and S42 constructs 23 and 24, with MD (potentially N-acetylated) and MK (never acetylated) N termini and the same -1 charge difference, gave essentially identical results, as did constructs 6 and 9 and constructs 10 and 15 (Fig. 2). These data are consistent only with the absence of N-acetylation in the S42 fusions prior to translocation.

Although we have not systematically tested the effect of distance from the TM segment on orientation signal strength, our data indicate that a single positive charge at -15 (constructs 21 and 22) has an effect similar to a negative charge at $+6$ (constructs 23 and 24), although charges separated by 31 or more residues (constructs 16 and 17) have no effect. The first charges in the α B reporter are 32 residues downstream of the TM segment, so they are presumably irrelevant to the charge difference signal. The first charges in the PB reporter are only 22 residues downstream but consist of the effectively neutral **KRSDTAE** cluster (Fig. 1) and are apparently also without effect on orientation.

The charge difference signal strength appears indifferent to local sequence context or to location relative to the TM segment, other than distance. This suggests a simple electrostatic response mechanism provided, in prokaryotes, principally by the transmembrane potential (10). However, there is no evidence for such a potential at the ER, although the interior and exterior differ sharply in redox potential and Ca^{2+} concentration. Recent studies in *E. coli* have demonstrated a role for anionic phospholipid concentration in preventing translocation of TM segments flanked by positive charges (12). Concentration of anionic phospholipids on the cytoplasmic face and depletion on the luminal face of the ER membrane could potentially provide the response mechanism to charge in eukaryotes. Such a distribution would be consistent with the observed distribution of anionic phospholipids in the plasma membrane of mammalian cells. This distribution is maintained by P-type ATPase lipid translocases, which have homologues in yeast (36). We are currently investigating yeast mutants with altered response to charged orientation signals and have identified mutants in a related gene.²

Effects of TM Sequence on Orientation—The charge-independent N_{exo} bias seen in all fusions containing the n20 Ste2p TM₁ segment is equivalent to a charge difference of about $+1$. Because the same bias was observed when the native Ste2p 51-residue N-terminal domain in S79 is reduced to 15 residues in S42, it seems unlikely that preferential translocation of this

³ Fred Sherman, personal communication.

N-terminal segment is a factor. Because the C-terminal PB and α B reporters also seem to be neutral, the TM₁ sequence itself is apparently responsible for this bias. Bias was essentially eliminated when the n20 TM₁ sequence was N-terminally truncated to about 15 residues or had its total hydrophobicity reduced to about 35 units on the GES scale (Fig. 7). Further reduction imposes an increasing N_{cyt} bias, equivalent to a charge difference of about -1 at 28 units, but insertion efficiency suffers, as seen in the S60-n14 and S60-n11 fusions. The bias is roughly equivalent to +1 at 42 units. Although such correlations are based on a small number of constructs, they are consistent with data on the effects of oligoleucine TM segments *in vitro* (25) and in mammalian cells (26), because a length of about 13 (36 units on the GES scale) is apparently neutral, and a length of 19 (54 units) overwhelms a -2 charge difference. Conservation of response mechanisms in eukaryotes is implied. This pattern is consistent with the normal function of secretion signals, which usually have a relatively short hydrophobic segment and a negative charge difference, both of which favor the required N_{cyt} insertion. A strong correlation was found between hydrophobicity peak height of secretion signals and SRP-dependence of secretion (15). By this criterion, translocation of even the n11 fusion should be SRP-dependent. There is no evidence, however, to suggest that utilization of the SRP-dependent or SRP-independent translocation mechanisms affect insertion orientation.

We constructed the S60-L₃-n11 and S60-L₅-n11 fusions in order to distinguish between the effects on orientation of TM segment length and total hydrophobicity. Although a better correlation with hydrophobicity was observed, these constructs introduced significant changes in hydrophobicity distribution, so potential effects on orientation must also be considered. Such effects were tested by moving the Leu pentapeptide from the N terminus in the S60-L₅-n11 fusion to the C terminus in the S60-n11-L₅ fusion. A dramatic effect on orientation was observed, suggesting that the more hydrophobic terminus of a TM segment is preferentially translocated. This was confirmed by analysis of the S60-HN and S60-HNrev fusions, which represent strong natural hydrophobicity gradients. TM segment sequence is clearly an important determinant of orientation, independent of total hydrophobicity, when the charge difference is small. Its maximum strength is about equal to that of a single charge difference. This novel observation introduces a new parameter to be incorporated into topology prediction algorithms. Application of such algorithms are planned and should provide tests of the universality of these observations.

Effects of TM sequence on orientation may extend beyond hydrophobicity. Aromatic residues, for example, are more frequent at the cytoplasmic boundary of TM segments, except for tryptophan residues, which are preferentially found at both boundaries (37). Such statistical biases in amino acid distribution have been successfully used in topology prediction (38). Responses to hydrophobicity gradients and to aromatic residues indicate a degree of sequence specificity, implying recognition by a protein, presumably some component of the translocon machinery. Because orientation appears to be determined by the sum of responses to independent signals, provided by TM

segment sequence and flanking charges, all must act in concert, presumably at the time of presentation to the translocon channel. Interaction of charged residues with a local electrostatic field, produced by membrane lipid head groups or possibly by a multimeric component of the core translocon, may combine with interaction of the hydrophobic core of the TM segment with hydrophobic protein domains of the translocon to determine the final orientation. A strong charge signal or an unusually lengthy TM segment can dominate the orientation signal. The extent to which natural observed TM protein orientations also result from selective destruction of inverted forms is currently unknown. The analysis of yeast mutants affecting responses to orientation signals should clarify these mechanisms. The importance of these responses is illustrated by the recent demonstration that neuropathology induced by the expression of prion protein mutants in mice and human correlates with membrane insertion of these highly conserved and unusual proteins in inverted orientation (5).

REFERENCES

- Hegde, R. S., and Lingappa, V. R. (1997) *Cell* **91**, 575-582
- Powers, T., and Walter, P. (1997) *Science* **278**, 2072-2073
- Siegel, V. (1997) *Cell* **90**, 5-8
- Dunlop, J., Jones, P. C., and Finbow, M. E. (1995) *EMBO J.* **14**, 3609-3616
- Hegde, R. S., Mastrianni, J. A., Scott, M. R., DeFea, K. A., Tremblay, P., Torchia, M., DeArmond, S. J., Prusiner, S. B., and Lingappa, V. R. (1998) *Science* **279**, 827-834
- von Heijne, G. (1990) *J. Membr. Biol.* **115**, 195-201
- Andersson, H., and von Heijne, G. (1993) *EMBO J.* **12**, 683-691
- Nilsson, I., and von Heijne, G. (1990) *Cell* **62**, 1135-1141
- von Heijne, G. (1992) *J. Mol. Biol.* **225**, 487-494
- Andersson, H., and von Heijne, G. (1994) *EMBO J.* **13**, 2267-2272
- Gafvelin, G., and von Heijne, G. (1994) *Cell* **77**, 401-412
- van Klompenburg, W., Nilsson, I., von Heijne, G., and de Kruijff, B. (1997) *EMBO J.* **16**, 4261-4266
- Kiefer, D., Hu, X. T., Dalbey, R., and Kuhn, A. (1997) *EMBO J.* **16**, 2197-2204
- Dalbey, R. E., Kuhn, A., and Heijne, G. v. (1995) *Trends Cell Biol.* **5**, 380-383
- Ng, D. T., Brown, J. D., and Walter, P. (1996) *J. Cell Biol.* **134**, 269-278
- Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 5786-5790
- Sato, T., Sakaguchi, M., Mihara, K., and Omura, T. (1990) *EMBO J.* **9**, 2391-2397
- Beltzer, J. P., Fiedler, K., Fuhrer, C., Geffen, I., Handschin, C., Wessels, H. P., and Spiess, M. (1991) *J. Biol. Chem.* **266**, 973-978
- Parks, G. D., and Lamb, R. A. (1991) *Cell* **64**, 777-787
- Parks, G. D., and Lamb, R. A. (1993) *J. Biol. Chem.* **268**, 19101-19109
- Denzer, A. J., Nabholz, C. E., and Spiess, M. (1995) *EMBO J.* **14**, 6311-6317
- Wallin, E., and von Heijne, G. (1995) *Protein Eng.* **8**, 693-698
- Cartwright, C. P., and Tipper, D. J. (1991) *Mol. Cell. Biol.* **11**, 2620-2628
- Harley, C. A., and Tipper, D. J. (1996) *J. Biol. Chem.* **271**, 24625-24633
- Sakaguchi, M., Tomiyoshi, R., Kuroiwa, T., Mihara, K., and Omura, T. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 16-19
- Wahlberg, J. M., and Spiess, M. (1997) *J. Cell Biol.* **137**, 555-562
- Komano, H., and Fuller, R. S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 10752-10756
- Chaudhuri, B., and Stephan, C. (1992) *FEBS Lett.* **304**, 41-45
- Wilson, C., Gilmore, R., and Morrison, T. (1990) *Mol. Cell. Biol.* **10**, 449-457
- Cartwright, C. P., Li, Y., Zhu, Y. S., Kang, Y. S., and Tipper, D. J. (1994) *Yeast* **10**, 497-508
- Redding, K., Holcomb, C., and Fuller, R. S. (1991) *J. Cell Biol.* **113**, 527-538
- Nacken, V., Achstetter, T., and Degryse, E. (1996) *Gene* **175**, 253-260
- Engelman, D. M., Steitz, T. A., and Goldman, A. (1986) *Annu. Rev. Biophys. Biophys. Chem.* **15**, 321-353
- Moerschell, R. P., Hosokawa, Y., Tsunasawa, S., and Sherman, F. (1990) *J. Biol. Chem.* **265**, 19638-19643
- Kulkarni, M. S., and Sherman, F. (1994) *J. Biol. Chem.* **269**, 13141-13147
- Tang, X., Halleck, M. S., Schlegel, R. A., and Williamson, P. (1996) *Science* **272**, 1495-1497
- Landolt-Marticores, C., Williams, K. A., Deber, C. M., and Reithmeier, R. A. (1993) *J. Mol. Biol.* **229**, 602-608
- Persson, B., and Argos, P. (1997) *J. Protein Chem.* **16**, 453-457

Transmembrane Protein Insertion Orientation in Yeast Depends on the Charge Difference across Transmembrane Segments, Their Total Hydrophobicity, and Its Distribution

Carol A. Harley, Jonathan A. Holt, Rhiannon Turner and Donald J. Tipper

J. Biol. Chem. 1998, 273:24963-24971.
doi: 10.1074/jbc.273.38.24963

Access the most updated version of this article at <http://www.jbc.org/content/273/38/24963>

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 38 references, 18 of which can be accessed free at <http://www.jbc.org/content/273/38/24963.full.html#ref-list-1>