The Role of Charged Residues in Determining Transmembrane Protein Insertion Orientation in Yeast*

(Received for publication, April 30, 1996)

Carol A. Harley and Donald J. Tipper‡
From the Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts 01655

The first 79 residues of the yeast Ste2p G protein-coupled pheromone receptor, including the negatively charged N-terminal domain, the first transmembrane segment, and the following positively charged cytoplasmic loop, has been fused to a Kex2p-cleavable β-lactamase reporter. Insertion orientation was determined by analysis of cell-associated and secreted β-lactamase activities and independently corroborated by analysis of membrane association and glycosylation patterns. This fusion inserts with exclusively N terminus exofacial (Nexo) topology, serving as a model type III membrane protein. Orientation is unaffected by removal of all three positively charged residues in the cytoplasmic loop or by deletion of all but 12 residues from the N-terminal domain. The residual —2 N-terminal charge apparently provides a signal sufficient to determine Nexo topology. This is entirely consistent with the statistically derived rule in which the charge difference, Δ(C–N), counted for the 15 immediately flanking residues, is the primary topology determinant. Mutations altering Δ(C–N) to zero favors Nexo insertion by 3 to 1, whereas increasingly negative values cause increasing inversion of orientation. All results are consistent with the charge difference rule and indicate that whereas positive charges promote cytoplasmic retention, negative charges promote translocation.

With the exception of the β-barrel proteins found in the outer membrane of Gram-negative bacteria, the vast majority of TM segments are α-helices of 18 or more mostly hydrophobic amino acid residues. Proteins of this type, whose mature forms span the membrane once, are commonly categorized into three major groups: Type I proteins contain a cleavable signal sequence at the N terminus. The signal recognition particle binds to this signal and targets the nascent polypeptide chain to the ER, where signal function initiates transfer of the mature N terminus of the protein across the ER membrane; a second hydrophobic domain acts as an anchor or stop-transfer sequence so that the mature protein adopts an Nexo topology. Type II proteins contain an uncleaved signal/anchor domain and adopt the opposite (Ncyt) topology. Type III proteins also contain an uncleaved anchor sequence, but the proteins insert with Nexo orientation. All three classes utilize the Sec machinery for insertion at the ER. Type IV proteins have a C-terminal signal/anchor domain and adopt the Ncyt topology but apparently use a different entry mechanism (3).

The topogenic signals that dictate the orientation of TM segment insertion consist principally of charged residues within the amino acid sequences flanking the TM segment (4). In prokaryotes, statistical analysis of sequences and extensive analyses of model proteins has led to formulation of the “positive inside rule”; positive charges impede translocation and cause retention of the adjacent end of the TM segment in the cytoplasm (4, 5). Arginine and lysine are of principal importance; histidine plays little part because of its low average degree of ionization at physiological pH. Negatively charged residues, aspartate and glutamate, also appear to have little effect. The dominance of positive charges is thought to reflect enhancement of ionization of amino groups and suppression of ionization of carboxylate groups on insertion into the membrane (6). This topogenic signal has no discernible sequence conservation, suggesting an electrostatic mechanism. The receptor for this signal is, at least in part, the TM potential at the bacterial cytoplasmic membrane (7, 8).

In eukaryotes, empirical analyses of the topology of TM proteins indicate a similar reliance of insertion orientation on adjacent charge, frequently interpreted as indicating compliance with the positive inside rule; in particular, analysis of the paramyxovirus HN protein, a model type II TM protein, has suggested a dominant effect for positively charged N-terminal residues (9–12). Orientation correlates best, however, with Δ(C–N), the difference in charges within an arbitrarily chosen window of 15 residues flanking the TM segment on either side (13). This “charge difference” rule differs from the positive inside rule in giving equal weight to positive and negative charges; the terminus with the most positively charged adjacent segment is retained in the cytoplasm so that positive and negative values of Δ(C–N) correlate with Nexo and Ncyt orientation, respectively. Statistically, a net charge difference of zero

* This research was supported by Council for Tobacco Research Grant 3777 (to D. J. T.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
‡ To whom correspondence should be addressed. Tel.: 508-856-2308; Fax: 508-856-5920; E-mail: dtipper@ummed.edu.
1 The abbreviations used are: TM, transmembrane; Ncyt, cytoplasmic N terminus; Nexo, exofacial N terminus; ER, endoplasmic reticulum; LSS, low speed supernatant; bp, base pair; PCR, polymerase chain reaction.
is also sequence to favor Nexo orientation (13). As in bacteria, the lack of sequence information in these topogenic signals implies an electrostatic mechanism. There is, however, no detectable potential across the ER membrane, implying the existence of some other electrostatic receptor. In addition to charged residues, there is a statistically significant difference in the content of aromatic amino acids, cysteine, and alanine between the exofacial and cytoplasmically disposed segments of integral membrane proteins, probably reflecting selection for function in different folding environments (14, 15).

Since insertion at the ER is cotranslational for most eukaryotic proteins, only the N-terminal domain preceding the first TM segment is potentially free to fold before translocation. In exofacial proteins, only the N-terminal domain preceding the first in different folding environments (14, 15).

Determinants of Membrane Protein Insertion Orientation

Materials and Methods

strains and reagents—All DNA manipulations were performed using the Escherichia coli strain DH5α (supE44 ΔlacU169 [F lacZΔM15] hsdS25 ΔlacZ75 ΔproAB thi-1 [F λpir]) [strain DH5α (supE44 ΔlacU169 [lacZΔM15] hsdS25 ΔlacZ75 ΔproAB thi-1 [F λpir]) were expressed in S. cerevisiae strain CRY1 (Mata ura3-1 leu2-3,112 trpl-1 his3-11 ade-2 can-1-100) (20). β-Lactamase was assayed using the synthetic substrate PADAC (7-thieryl-2-acetamido)-3-[2-(4-N-Dimethylaminothiazole)] pyridiniummethyl-3-epicholyl-carboxylic acid] (Calbiochem), as described previously (21). Renografin-76 (Squibb) was used for density gradient fractionation procedures (22). All PCR amplifications were carried out using Pfu DNA polymerase (Stratagene Cloning Systems).

Expression Vectors—Fusion constructs are designated as either PB, αB, or SPB derivatives. P is a processing fragment from the K. prepoxin encoding one internal and one C-terminal Kex2p cleavage site and three N-glycosylation sites (20). α is a fragment of prerepro-α-factor containing one C-terminal Kex2p cleavage site and two N-glycosylation sites (20). K, prepoxin sequence and two sequence of β-lactamase (23). In p21-PGK-SPB, SPB is inserted between the PGK promoter and transcription terminator fragments in the multiple cloning site of the episomal URA3 vector Yep352 (21). Expression results in translocation of PB; following cleavage of PB by Kex2p, mature β-lactamase is secreted, preceded by the tripeptide Asp-Pro-Gly; pGAL-SPB is identical to p21-PGK-SPB, except that the PGK promoter is replaced by the galactose inducible GAL1I promoter (21). p32-SPB is derived from p21-PGK-SPB by elimination of the Pet site in B, the βGI and SalI sites flanking the PGK terminator, the vector SalI sites, and the BamHI and BstEI sites between P and B. Following cleavage by Kex2p, mature β-lactamase is secreted, preceded by the tripeptide Asp-Leu-Arg; expression is unaffected. The Pet site is derived by insertion of a unique NolI site in place of the XhoI site at the N terminus of S (AT GCA TCC ATO) and restoration of the SalI site following the PGK terminator. The NolI site is in the same reading frame as the unique Pet site at the N terminus of P, facilitating interconversion of SPB and PB fusions.

DNA Manipulations—The first 250 bp of STE2, encoding the first 79 amino acid residues of Ste2p, were amplified by PCR using the N- and C-terminal oligonucleotide primers 5ʹ GGCGTCAAGTAGCTGCTGATTGGAGGGCGCCGC 3ʹ (where K = G and T = G or C) A mixture pool of full length 250-bp PCR products were obtained and cloned as Xhol/PstI fragments into the corresponding sites of pS79a-PB. The resulting constructs were transfected into yeast, and the stable transformants were isolated on plates containing 2% glucose. The following C-terminal Ste2p mutants, all of which contain the NdeI site marking the R58I mutation, were identified by sequence analysis: S79b, S79c, S79d, and S79e, containing the mutations R74T, R76T, K77T, and R74T/R76T respectively. The 5ʹ GGCGTCAAGTAGCCTGCTGATTGGAGGGCGCCGC 3ʹ (where K = G and T = G or C) using the mutagenic primers 5ʹ ATGTTTTGGTGTCAATGTTGGGCAGC 3ʹ, which create an NdeI site and a single amino acid mutation, R58I in the TM region of Ste2p, and 5ʹ GACATTCTTTCATGATGGCGACAC 3ʹ, which remove the NolI site within the URA3 gene. The product is the pS79a-PB construct. All other mutants are derived from pS79a-PB; sequences of all mutants were confirmed.

Construction of C-terminal Mutants (URA3 Derivatives)—Using pS79a-PB as a template, the previously described N-terminal STE2 oligonucleotide primer (Xho site) and the degenerate C-terminal primer 5ʹ GGCGTCAAGTAGCCTGCTGATTGGAGGGCGCCGC 3ʹ (where K = G and T = G or C) using the mutagenic primers 5ʹ ATGTTTTGGTGTCAATGTTGGGCAGC 3ʹ, which create an NdeI site and a single amino acid mutation, R58I in the TM region of Ste2p, and 5ʹ GACATTCTTTCATGATGGCGACAC 3ʹ, which remove the NolI site within the URA3 gene. The product is the pS79a-PB construct. All other mutants are derived from pS79a-PB; sequences of all mutants were confirmed.

Construction of αB Derivatives—The P fragment of constructs pS79a, S79b, S79c, and S79d-PB was replaced by a fragment of prepro-α-factor with a C-terminal Kex2p processing site, to eliminate the potential contribution of charges within P to topological signals. The fragment was amplified by PCR using the oligonucleotide primers 5ʹ GGCGTCAAGTAGCCTGCTGATTGGAGGGCGCCGC 3ʹ (where K = G and T = G or C) using the mutagenic primers 5ʹ ATGTTTTGGTGTCAATGTTGGGCAGC 3ʹ, which create an NdeI site and a single amino acid mutation, R58I in the TM region of Ste2p, and 5ʹ GACATTCTTTCATGATGGCGACAC 3ʹ, which remove the NolI site within the URA3 gene. The product is the pS79a-PB construct. All other mutants are derived from pS79a-PB; sequences of all mutants were confirmed.

Construction of N-terminal and Double Mutants (LEU2 Derivatives)—To facilitate cloning, LEU2 derivatives of both pS79a-PB and pS79a-PB were constructed as follows: the LEU2 yeast expression vector Yep351 was digested with HindIII/VspI. The 3.4-kbp fragment containing the LEU2 selectable marker and 2 micron DNA was cloned into the corresponding sites in pS79a-PB and pS79g-PB, producing pS79a-PB-L and pS79g-PB-L. Using pS79a-PB-L as a template and the degenerate C-terminal oligonucleotide primer 5ʹ GGCGTCAAGTAGCCTGCTGATTGGAGGGCGCCGC 3ʹ (where K = G and T = G or C), the C-terminal oligonucleotide primer 5ʹ CCAGCTGACGGCATACCT 3ʹ, a mixture of N-terminally truncated 120-bp PCR products were obtained and cloned as Xhol/PstI fragments into the corresponding sites of pS79a-PB-L. In these mutants the
Density Gradient Fractionation—Cells were grown at 30 °C in SC (synthetic complete) Ura− or Leu− medium buffered with 50 mM BisTris (pH 7.0) and harvested in early postexponential growth phase (2 × 10^7 to 3 × 10^7 cells/ml), and the supernatant was removed and the pellet fractions (P100) were resuspended in Buffer A (20 mM Tris-HCl, pH 7.6, 50 mM NaCl, 20 mM Na_2PO_4, 10 mM KF, 2 mM EDTA) containing 0.2 mM phenylmethylsulfonyl fluoride and 2 mg/ml pepstatin A and resuspended in the same buffer. The cells were broken by vortexing in the presence of glass beads (0.45 μm), and large cell debris was removed by centrifugation at 330 × g for 5 min, producing a low speed supernatant (LSS). The β-lactamase activities in culture supernatants and LSS extracts were determined spectrophotometrically using a synthetic substrate, PADAC, and a microtiter plate assay, as described previously (21).

Cell Fractionation and β-Lactamase Activity Assay—Cells were grown at 30 °C in SC Ura− or Leu− medium and harvested in early postexponential growth (2–3 × 10^7 cells per ml). Cells (2 ml) were then washed once with 1 ml of Buffer A, once with 1 ml of TE buffer (50 mM Tris-HCl, 1 mM EDTA, pH 7.5; containing the same protease inhibitors) and then suspended in 1 ml of TE buffer. The cells were broken by vortexing in the presence of glass beads, and the LSS was isolated as described above. 0.5 ml of this LSS was mixed with either 10% Triton X-100 or 1 ml Na_2CO_3, pH 12.4 (9.1 by volume; final concentrations, 1% and 0.1 M, respectively). Mixtures were incubated for 20 min on ice and then centrifuged at 14,000 × g for 30 min, producing a clear supernatant (S100) and an insoluble pellet (P100). S100 (1 ml) was centrifuged for 10 min at 20,000 × g to remove any remaining insoluble material. The clear supernatant (S100) was then diluted 1:10 with Buffer A. A total of 300 μl of this diluted supernatant was then mixed with 300 μl of Buffer A. The mixture was vortexed and then centrifuged at 14,000 × g for 20 min.

RESULTS

We previously confirmed (19) that the Ste2p receptor adopts the predicted seven TM segment topology common to G protein-coupled receptors, with an exofacially exposed N terminus and a cytoplasmically disposed C terminus (Fig. 1A). Analysis involved the in vivo expression of C-terminal fusions of Ste2p fragments, starting after the second TM segment, to a reporter, P. Encodes a fragment of the K_2 killer preprotoxin that contains two sites efficiently cleaved by the Kex2p protease (23); B, the mature form of β-lactamase, is downstream of the second Kex2p site (Fig. 2A). Kex2p is located in the late Golgi with its active site in the lumen (24) so that cleavage of the PB reporter and secretion of B occurs only if the fusion site is luminal (exofacial). Secreted β-lactamase activity accumulates stably (half-life, about 8 h) in media buffered at pH 6.5–7 (19). Analysis of the ratio of secreted to cell-associated β-lactamase activity, therefore, can be used to determine whether the fusion site is cytoplasmic or exofacial. Exofacial location of the Ste2p-PB fusion was demonstrated in this manner (Fig. 1B; Ref. 5).

The reliability of interpretation of this data depends first on the relative stabilities of secreted β-lactamase (from N_exo fusions) and of N_cyt fusions where the PB reporter is cytoplasmic and unprocessed; second, on the kinetics with which fusions with an exofacial PB reporter are translocated to the trans-Golgi and processed by Kex2p; and third, on the efficiency with which processed β-lactamase is released into the medium. Reliable correction factors for these factors are now provided by analysis of the fate of control secreted and N_cyt PB fusions. Although the specific activities of the various fusions have not been determined directly, signal strengths in Western blots in these and previous studies (19) are always consistent with measured activity, when compared to that of a purified β-lactamase standard.
that secretion and Kex2p processing is effectively complete. The data in Fig. 4 are corrected for the 20% underestimate of secreted activity represented by supernatant activity.

The control construct cPB is essentially identical to p34nSPB except that S is replaced by the first 8 residues of prepro-α-factor. All β-lactamase activity from cPB expression is cell-associated and is found in the S100 supernatant from Airfuge fractionation of broken cells and so is cytoplasmic. Activity accumulates to the same level as secreted activity in cells expressing p34nSPB (Table I), so cytoplasmic β-lactamase has similar stability.

The S79-PB Fusion Is Inserted Exclusively with Nexo Orientation—Although exofacial location of the Ste29-PB fusion site implies that TM1, the first transmembrane segment of Ste2p, inserts into the bilayer with an Nexo orientation, this was not independently demonstrated (19). We constructed the S79-PB fusion for this purpose. The S79 fragment comprises the first 79 residues of Ste2p, which include an N-terminal fragment of approximately 50 residues, TM1, and the 8-residue first cytoplasmic loop (Fig. 2A). This loop (TSRSRKTP) carries three positive charges, predicted to favor Nexo insertion of the S79-PB fusion (Fig. 1C). If insertion is efficient, it should provide an easily manipulable model for the insertion of the first TM region of G protein-coupled receptors (15) and of Type III (Nexo)-oriented proteins in general.

P has three N-glycosylation sites (Fig. 2A), and in secreted PB, all three are core-glycosylated in the ER; elongation in the Golgi occurs, but is obvious only in the absence of Kex2p function (23). Full-length Ste2p has four potential N-glycosylation sites but only two of these, at Asn25 and Asn32, in the exofacial N terminus, lie sufficiently distant from TM regions to be accessible to the glycosylation machinery (25). The predominant species seen in Western blots are consistent with core-glycosylation at either one or both sites (22). In Nexo-inserted S79-PB, therefore, not only should the reporter be cytoplasmically disposed, giving exclusively cell-associated β-lactamase activity, but processing of P should be prevented, leaving an intact, membrane-associated fusion protein with a glycosylated Ste2p N terminus (Fig. 2A). Any fraction inserted in the opposite Ncyt orientation will be glycosylated on P and, if also translocated to the Golgi, will be processed by Kex2p to produce secreted, mature β-lactamase. Orientation, therefore, can be confirmed by Western blot analysis to detect the intact fusion protein and to determine its state of glycosylation. Size predictions are given in Table I.

Expression of the S79-PB fusion produced only cell-associa-
membrane fraction, and these were reduced to a single band of 44 kDa by treatment with endoglycosidase H (Fig. 5, C, lanes 1 and 2), consistent with the deduced glycosylation pattern. The 30-kDa species seen only after endoglycosidase H treatment and detected only with anti-β-lactamase (Fig. 5, lane C) comigrated with mature β-lactamase (Fig. 5, lane 9) and probably results from artifactual cleavage at the Kex2p sites during the prolonged incubation with endoglycosidase H. The corresponding 8.7-kDa Ste2 fragment is too small to be retained on the gel shown in Fig. 5D. The Ste79-PB fusion, therefore, is efficiently inserted into the membrane in an Nexo orientation and appears to be a mixture of species glycosylated on either one or both of the sites in the Ste2p N terminus.

Membrane association of the fusion protein was also demonstrated by Renografin gradient fractionation of disrupted cells. Following fractionation, three bands were distinguished using markers identified by specific antisera: low density membranes, corresponding to vacuole, Golgi, and ER, are marked by the Ost1p (Fig. 6, A) and negative charges in the C-terminal RSRK motif; C, αβ fusions; D, S42 fusions retaining the RSRK motif; E, S42 fusions lacking the RSRK motif.

Because failure in the processing of S79-PB by Kex2p could reflect failure in translocation to the late Golgi, expression was repeated in the presence of ER-Kex2p. ER-Kex2p is a derivative of Kex2p lacking the C-terminal TM region, which is responsible for its normal localization in the late Golgi; instead, it has a C-terminal HDEL sequence that causes it to locate primarily in the lumen of the ER, where it becomes active (26). When coexpressed with ER-Kex2p, S79-PB still produced only cell-associated activity (data not shown). None of this activity, therefore, appears to result from protein-inserted Nexo that has failed to translocate to the Golgi. In conclusion, all of the detected S79-PB fusion protein is a TM protein, inserted into the membrane in an Nexo orientation, and appears to be a mixture of species glycosylated on either one or both of the sites in the Ste2p N terminus.

Membrane association of the fusion protein was also demonstrated by Renografin gradient fractionation of disrupted cells. Following fractionation, three bands were distinguished using markers identified by specific antisera: low density membranes, corresponding to vacuole, Golgi, and ER, are marked by the Ost1p (Fig. 6, A) and negative charges in the C-terminal RSRK motif; C, αβ fusions; D, S42 fusions retaining the RSRK motif; E, S42 fusions lacking the RSRK motif.

Because failure in the processing of S79-PB by Kex2p could reflect failure in translocation to the late Golgi, expression was repeated in the presence of ER-Kex2p. ER-Kex2p is a derivative of Kex2p lacking the C-terminal TM region, which is responsible for its normal localization in the late Golgi; instead, it has a C-terminal HDEL sequence that causes it to locate primarily in the lumen of the ER, where it becomes active (26). When coexpressed with ER-Kex2p, S79-PB still produced only cell-associated activity (data not shown). None of this activity, therefore, appears to result from protein-inserted Nexo that has failed to translocate to the Golgi. In conclusion, all of the detected S79-PB fusion protein is a TM protein, inserted into the ER exclusively in an Nexo orientation.

The S79-PB Fusion Is Also Inserted Exclusively with Nexo Orientation—The topology of insertion of the Ste2p fragment was further analyzed in the S79-PB construct in which the S secretion signal is placed between the fusion site and the PB terminator. The topology of Ste2p is controversial (27, 28; see “Discussion”). The S79-PB fusion, therefore, is inserted into the membrane in an Nexo orientation and appears to be a mixture of species glycosylated on either one or both of the sites in the Ste2p N terminus.

The increases in size due to core glycosylation are estimates.

### TABLE I

| Peptide fragment | Residues | kDa | Nexo (%) | After endo H (%) |
|------------------|----------|-----|----------|-----------------|
| S79, S79a | 81 | 8.7 | 15 | 9 |
| S42 | 44 | 4.8 | 33 | 8 |
| S | 35 | 3.7 | 59 | 6.4 |
| P | 59 | 6.4 | 15 | 7 |
| α | 31 | 3.5 | 9 | 4 |
| B | 266 | 29.3 |
| S79-PB | 44 | 50 | 44.4 |
| PB | 35.5 | 44 | 36 |
| S79-S | 42.2 | 19 |
| S42-PB | 40.5 |
| S42-PB | 40.5 | 49 | 41 |

### Determinants of Membrane Protein Insertion Orientation

**Fig. 4. β-Lactamase activity data for the indicated control constructs and fusions.** Secreted activity, Asec, and cell-associated activity, Aaca (ng of β-lactamase/107 cells), are corrected for the lag in secretion and for the turnover of cell-associated fusion protein using the formula: cytoplasmic PB/luminal PB = Aaca/Asec × B.PGK/GAL1. See “Results.” All data shown are averages from at least two independent transformants, each analyzed at least three times. Δ(C–N) is the charge difference calculated according to Hartmann et al. (13). The charged residues contributing to this signal that flank the TM segment (box) are indicated in single-letter code with positive charges in boldface and negative charges in outline font. The total and secreted activity, Asec, and cell-associated activity (data not shown) were corrected for the lag in secretion and for the turnover of cell-associated fusion protein using the formula: cytoplasmic PB/luminal PB = Aaca/Asec × B.PGK/GAL1. See “Results.” All data shown are averages from at least two independent transformants, each analyzed at least three times. Δ(C–N) is the charge difference calculated according to Hartmann et al. (13). The charged residues contributing to this signal that flank the TM segment (box) are indicated in single-letter code with positive charges in boldface and negative charges in outline font. The total and secreted activity, Asec, and cell-associated activity (data not shown) were corrected for the lag in secretion and for the turnover of cell-associated fusion protein using the formula: cytoplasmic PB/luminal PB = Aaca/Asec × B.PGK/GAL1.
Fig. 5. Western blot analysis of cell-associated fusion proteins from cells of strain CRY1 expressing the indicated plasmid constructs. Unless otherwise stated, the primary antibody used in detection was anti-β-lactamase. A, S79-SPB and S79a-, S79f-, and S79g-PB fusions. Supernatant (S) and pellet (P) fractions from Airfuge separation of proteins in the LSS from broken cells were analyzed by SDS-polyacrylamide gel electrophoresis. The LSS was either untreated (Ctrl), solubilized with 1% Triton X-100 (Trit), or incubated with 0.1 M Na2CO3 (Carb) prior to fractionation. B, S42o-, S42p-, S42u-, and S42t-PB fusions; fractionation as in A. C, Triton-solubilized LSS proteins from cells expressing the indicated fusions were analyzed with (+) and without pretreatment with endoglycosidase H. D, some of the same protein samples, with (+) and without pretreatment with endoglycosidase H, detected with antibody to the Ste2p N terminus.

Fig. 6. Renografin gradient fractionations of proteins in the low speed supernatant from broken cells. Fractions (lane 1, top) were analyzed by Western blot using antisera to the Pma1p plasma membrane ATPase and the Ost1p Golgi oligosaccharoyl transferase as markers. Fusion proteins were detected using anti-β-lactamase. The cytoplasmic fractions were identified using cytoplasmically expressed β-lactamase in a separate gradient. A, marker proteins and the S79-PB fusion. B, marker proteins and the S79a-PB fusion.

promoter is shown), presumably reflecting lower messenger stability or translational efficiency in the Ste2p fusion, but showing that expression levels are comparable and much higher than in fusions to longer Ste2p fragments (19). When cells expressing the GAL-driven fusion were shifted from galactose (inducing) to glucose (repressing) medium, an additional 10% of the total activity was chased into the culture supernatant; the remaining 5–10% was solubilized on spheroplast formation. The activity that is cell-associated at steady state is, therefore, in transit within the secretory pathway or trapped in the cell wall, indicating that all of the PB reporter in the S79-SPB fusion is translocated into the ER lumen. This was confirmed by Western blot analysis of the transiently cell-associated material. Anti-β-lactamase serum detected products only in the S100 fraction (Fig. 5A, lanes 7–12). Since the intact SPB fusion would have been readily detected in the membrane fraction, efficient translocation and cotranslational cleavage by signal peptidase at the S-PB junction is indicated. Signal function in this fusion, therefore, is not compromised by the preceding S79 fragment. The major species of about 48 kDa collapsed to 40 kDa on treatment with endoglycosidase H (Fig. 5C, lanes 3 and 4), as predicted for luminal, soluble, core-glycosylated PB (Table I). Mature β-lactamase (30 kDa) is seen only after endoglycosidase H treatment and probably results from artificial cleavage at the Kex2p sites during incubation with endoglycosidase H.

Although translocation of the PB reporter is consistent with Nexo insertion of the Ste2 fragment, with S acting as an efficient signal sequence (Fig. 1E), two alternate insertion events could lead to PB translocation. First, S could perform its normal secretion signal function while the Ste2 fragment failed to insert. Second, we previously showed that S functions poorly as a stop-transfer segment (19), so PB would remain luminal even if the fusion inserted Ncyt. These possibilities were also distinguished by Western blot. If translocation of the S79 N terminus is efficient, the S79-S fragment released by signal peptidase action should remain in the membrane fraction; antiserum to the Ste2p N terminus detected products only in the membrane fraction (Fig. 5D, lanes 3 and 4). No signal corresponding to the intact fusion was visible; two species of apparent size 19 and 16 kDa were seen, and these collapsed to a single species of 13 kDa after treatment with endoglycosidase H. These are the sizes predicted for the S79-S fragment carrying either two or one core N-CHO fragments and their endoglycosidase products (Table I). These results confirm efficient Nexo insertion of the Ste2 N terminus and efficient cleavage by signal peptidase in this S79-SPB fusion (Fig. 1E) to give the secreted PB product and a transmembrane S79-S fragment. A clear indication of the core glycosylation pattern in the Ste2p N-terminal fragment is also provided, consistent with mobilities seen in the S79-PB fusions and presumably reflecting that seen in the intact receptor (22).

Relation of the Distribution of β-Lactamase Activities to the Nexo/Ncyt Ratio—As shown above, culture supernatant β-lac-
Determinants of Membrane Protein Insertion Orientation

24631

tamase activity, $A_{sec}$, underestimates translocated PB from both p34nSPB and S79-SPB by about 20%. $A_{sec}$, therefore, is multiplied by 1.25 to estimate the total secreted fraction due to luminal PB. The half-life of the cell-associated S79-SPB fusion, determined by expression from the GAL1 promoter and shift from galactose to glucose medium, was only 2 h compared to the 8 h for $A_{sec}$. As a consequence, accumulated cell-associated $\beta$-lactamase activity ($A_{ca}$) from S79-SPB was 4-fold lower than $A_{sec}$ from the S79-SPB fusion. To calculate the true ratio of cytoplasmic and luminal reporter in PB fusions, therefore, we used the following formula: cytoplasmic PB/luminal PB = $N_{exo}$/ $N_{cyt}$ = 1.25 $A_{ca}/A_{sec}$ (Fig. 4). Because of these factors, measured activities are very sensitive to a small fraction of luminal reporter when the reporter is predominantly cytoplasmic ($N_{exo}$ insertion) but are insensitive to a small fraction of cytoplasmic reporter when the reporter is predominantly luminal ($N_{cyt}$ insertion).

Elimination of Arg\textsuperscript{26} in Ste2 TM\textsubscript{1} Affects neither Stability nor Orientation of the PB Fusion—Charged amino acid residues within TM segments of integral membrane proteins, particularly when near the center, can result in the retention and degradation of these proteins within the ER (27). We therefore tested the effect of replacing Arg\textsuperscript{26} in the S79-SPB fusion, near the center of Ste2p TM\textsubscript{1}, with Ile\textsuperscript{28}, producing S79a-SPB (Fig. 2A). Neither stability nor orientation were affected (Fig. 4B), suggesting that Arg\textsuperscript{26} does not destabilize this fusion in yeast. Western blot analysis confirmed that as for the S79-SPB fusion, all of the activity was cell-associated and all was in the form of an integral membrane protein of unaltered size and apparent molecular mass (Fig. 5A, lanes 13–18), as confirmed by the results of endoglycosidase H treatment and detection with either anti-$\beta$-lactamase or anti-Ste2p (Fig. 5, C and D, lanes 5 and 6). We used the R581 mutant to construct all subsequent fusions to eliminate any possible effects of the Arg\textsuperscript{26} charge on TM\textsubscript{1} insertion orientation.

Neither Elimination of the Three Positive Charges following TM\textsubscript{1} nor Increasing the Separation Between TM\textsubscript{1} and the First Charges of the Reporter Affects Orientation of the PB Fusion—Apart from the positively charged N-terminal methionine, the N-terminal $N_{exo}$ domain of the Ste2\textsubscript{2} fragment contains only negative charges, and only two of these, Asp\textsuperscript{39} and Glu\textsuperscript{40}, fall within the arbitrary 15-residue window to either side of TM\textsubscript{1} used by Hartmann et al. (13) in calculating the charge difference, $\Delta$(C–N) (Fig. 2A). Since the first cytoplasmic loop following TM\textsubscript{1} contains three positive charges in the RSRK motif, $\Delta$(C–N) for S79-SPB and S79a-SPB is +5 (Fig. 4B). To test the importance of the positively charged residues in the RSRK motif, they were mutated, individually and in combination, to T (constructs S79b-SPB to S79g-SPB). All of these fusions had the same properties and data for three representative constructs, in which $\Delta$(C–N) is reduced to +4, +3, or +2, are presented in Fig. 4B. In each case, the expression level was essentially the same as in S79a-SPB, and essentially all of the activity was cell-associated. All of these cell-associated materials behaved as integral membrane proteins, as shown for S79c-SPB and S79g-SPB ($\Delta$(C–N) = +2 and +3, respectively) in Fig. 5A, lanes 16 and 17, respectively. The major species indicated a glycosylation pattern unchanged from the unpurified PB fraction, and all data imply the same exclusive $N_{exo}$ insertion orientation. Renografin fractionation of disrupted cells expressing the most extreme mutant, S79g-SPB, showed essentially the same pattern as S79-SPB (Fig. 6B); the activity cosedimented with the Ost1p membrane marker. It is clear, therefore, that the charges in the RSRK motif of S79-SPB are a dispensable part of the topogenic signal.

Although these data suggest that the N-terminal negative charges in the S79g-SPB fusion may be sufficient to determine insertion orientation, it remained possible that the receptor for this signal senses charges located further from TM\textsubscript{1} within the PB reporter. The closest charges are found in the effectively neutral KRSDTAE motif of P, starting 21 residues after TM\textsubscript{1} (Fig. 2A). To test its potential role, P was replaced with $\alpha$, a fragment of prepro-$\alpha$-factor also terminating in a Kex2p processing site but in which the first charges, present in the KEE motif, are separated from TM\textsubscript{1} by 32 residues (Fig. 2B). Expression of these four representative $\alpha$B fusions, S79a-AB to S79g-AB (Fig. 3D), produced exactly the same results as their PB versions; all of the activity remained cell-associated, consistent with an $N_{exo}$ orientation (Fig. 4C). Expression levels were higher, suggesting a modest increase in the in vivo stability of these fusions. All of the fusion proteins were integral membrane proteins (data not shown). It appears that insertion remains efficient in these fusions and that orientation is unaffected by the location and nature of the proximal charges in the PB and $\alpha$B reporters.

Elimination of the Charge Difference Leads to Partial Inversion, Exaggerated by Charge Inversion—The results described so far are entirely consistent with the charge difference rule, if equal weight is given to positive and negative charges; in addition, the results with S79g-AB and S79g-AB indicate that the residual topogenic signal in these fusions promotes translocation and $N_{exo}$ insertion with high efficiency. According to the charge difference rule, this signal is the two negative charges on Asp\textsuperscript{39} and Glu\textsuperscript{40}, located about 10 residues upstream of TM\textsubscript{1}. However, it remained possible that insertion orientation in these fusions was affected by the more distant negative charges at Asp\textsuperscript{14} and Asp\textsuperscript{20} or reflected some unique property of the Ste2 N-terminal sequence in addition to its negative charges. To resolve this issue, we replaced Ste2p residues 1–38 with Met-Asp, producing a truncated, 42-residue Ste2 fragment called S42 (Figs. 2A and 3E) with an MDDE N-terminus. Since the positively charged N-terminal Met is now close to TM\textsubscript{1}, this construct has the same effective charge difference of +5 as S79a-AB. The N-terminal domain upstream of TM\textsubscript{1} is truncated to 12 residues and should no longer be N-glycosylated. We constructed a series of S42-AB fusions that contained the three positive charges in the C-terminal RSRK motif and in which the MDDE motif was mutated by PCR to provide a net N-terminal charge that ranged from −2 to +4. The fusions, in order of decreasing charge difference, are called S42h-AB to S42n-AB (Fig. 4D).

In the S42h-AB to S42k-AB series, in which $\Delta$(C–N) = 5, 4, or 2, all of the $\beta$-lactamase activity remained cell-associated (Fig. 4D) and all of this activity was in the form of an unglycosylated integral membrane protein of the predicted size (41 kDa; Table I; data not shown). In S42l-AB, in which $\Delta$(C–N) = +1, about 8% inversion was observed; this was increased to 26 and 47%, respectively, in S42m-AB and S42n-AB, in which $\Delta$(C–N) = 0 and −1, respectively (Fig. 4D). In these fusions, TM\textsubscript{1} is flanked by two groups of high positive charge; however, all of the cell-associated activity from these fusions remained in the P100 fraction, showing that this high charge density did not impede membrane insertion (data not shown). Thus, whereas a charge difference of +2 provided a strong topogenic signal and near exclusive $N_{exo}$ orientation, a unit charge differences was more ambivalent, allowing some inversion. There appears to be a bias toward $N_{exo}$ insertion in these constructs, as this still predominated when $\Delta$(C–N) = 0. Reversal of the unit charge difference lead to a 50:50 ratio of insertion orientation, also consistent with a bias toward $N_{exo}$ insertion.

The predominant charges in these fusions are positive; replacing the C-terminal RSRK motif in S42h-AB to S42n-AB
with the neutral TSTT peptide from S79g produced the S42o- to S42u-PB fusions (Fig. 4E). The topogenic signal in S42o-PB and S42p-PB, in which Δ(C–N) = +2 and +1, respectively, results entirely from a net N-terminal negative charge, as in S79g-PB. Activity data indicates 92–94% Nexo insertion (Fig. 4E), and this is entirely in the form of an unglycosylated integral membrane protein (Fig. 5B, lanes 1–12) of the expected 41-kDa size, unaffected by treatment with endoglycosidase H (Fig. 5C, lanes 7 and 8). This emphasizes the ability of negative charges, even in the context of the truncated N-terminal domain, to promote translocation of the TM, N terminus. A Δ(C–N) of −1 due to the single net N-terminal positive charge in S42q-PB and S42r-PB caused reversal of insertion, more completely than in the highly charged S42n-PB (Fig. 4E). Addition of one, two, or three additional N-terminal positive charges led to near total inversion in S42s-PB and complete inversion in S42t-PB and S42u-PB (Fig. 4E). Although orientation appeared to be exclusively Ncyt, our assay procedure would be insensitive to 5% or less Nexo insertion. Complete inversion of orientation was, however, confirmed by Western blot analysis.

The S42s-PB, S42t-PB, and S42u-PB β-lactamase fusion proteins were located in the high speed pellet (P100 membrane) fractions from broken cells, virtually all as integral membrane proteins, solubilized by 1% Triton but not by 0.1 M Na2CO3 (Fig. 5B, lanes 13–30). Most of the S42s-PB fusion and all of the S42t-PB and S42u-PB fusions migrated on SDS-polyacrylamide gel electrophoresis as a pair of 47–50-kDa proteins that collapse to a species of about 42 kDa on treatment with endoglycosidase H (Fig. 5C, lanes 10–13). These species have the sizes predicted for the Ncyt-inserted fusions, core N-glycosylated at two and three of the sites in the PB reporter (Table I), and must represent translocated fusion protein in transit to the Golgi prior to Kex2p cleavage. No signal is seen for mature β-lactamase in the sample prior to endoglycosidase H treatment, implying rapid kinetics for export following Kex2p cleavage. The β-lactamase seen after endoglycosidase H treatment is presumably an artifact of proteolysis. Only in S42s-PB is an unglycosylated 41-kDa species corresponding to Nexo-inserted fusion visible (lanes 13–18); it accounts for about 15% of the total signal, consistent with the activity data (Fig. 4E). The absence of this species in the S42t-PB and S42u-PB fusion products confirms their exclusive Ncyt insertion.

These data show that the S79 fragment and its truncated S42 form have redundant topogenic signals: either the C-terminal positive charges or the N-terminal negative charges being sufficient to determine insertion in Nexo orientation, strictly in accordance with the charge difference rule.

DISCUSSION

S79, the 79-residue N-terminal fragment of Ste2p, includes the first transmembrane segment and flanking charged topogenic sequences of this receptor. Fusion to the PB or αB Kex2p-cleavable β-lactamase reporters produced model Type III membrane proteins that inserted in vivo in the Nexo orientation native to the Ste2p N terminus and provided an assay system for quantitatively monitoring this insertional orientation. We have exploited this assay to test the effect of changes in flanking charge on orientation, allowing us to draw several conclusions concerning topogenic signals for eukaryotic membrane protein insertion, an issue that has previously been addressed only to a limited extent, using principally in vitro systems that may not fully reflect the intracellular environment. It was necessary, however, to first establish the validity of the assay technique.

In this fusion technique, the ratio of cytoplasmic and exofacial fusion sites for the PB reporter is deduced from the ratio of cell-associated and secreted β-lactamase activities. For integral transmembrane protein fusions, this ratio translates directly into the ratio of Nexo and Ncyt insertion orientation. Validity, therefore, depends on a demonstration that all of the cell-associated activity is indeed in the form of integral membrane proteins and on allowance for incomplete release of secreted β-lactamase into the medium and any differences in specific activity and turnover rates of cell-associated fusion proteins and secreted β-lactamase. Through analysis of appropriate controls, we have shown that all of the cell-associated fusions are integral membrane proteins, that culture supernatant activity underestimates secretion by 15–20%, that PB fusion proteins and free β-lactamase have similar specific activities, and that correction is needed for the 4-fold shorter half-life of membrane-associated PB fusion proteins. As a consequence, the confidence limits for detection of Ncyt insertion in a construct that gives near exclusive Nexo insertion are ±2%, whereas those for Nexo insertion in a construct that gives near exclusive Ncyt insertion are only ±8%. Western blot analyses of glycosylation patterns, however, provide an independent test of insertion orientation, and results with both the S79-PB and S79-SPB fusions were entirely consistent with the exclusive Nexo topologies deduced from activity distribution data. Besides confirming translocation of the N terminus of the Ste2p fragment, the glycosylation pattern of the SPB fusion confirmed the suspected modification pattern in the parent Ste2p receptor (22). We can conclude, therefore, that at least 98% of both fusions insert in Nexo orientation. These observations effectively complete our previous in vivo topological analysis of Ste2p (19). The only remaining issue is bias introduced by the reporters. First, all components of the PB and αB reporters are normally secreted by yeast with high efficiency (23), so they are not likely to impede translocation. Second, reporter neutrality is implied by the identical orientations seen for PB and αB fusions, in which the separation of the TM segment from charged residues in the reporter is increased from 21 to 32 residues. Finally, neutrality is also implied by the inversion of orientation seen when the proximal flanking charge is reversed.

If the PB reporter is neutral, then the S79 fragment must contain all of the topogenic information necessary for determining its Nexo orientation. It has been clear for some time that the major topogenic signal for determining TM segment insertion orientation is provided by flanking charged residues. In bacteria, the validity of the positive inside rule codifying the predominant role of positively charged residues in this signal has been firmly established (4, 5). The relative suppression of the effects of negative charges has been attributed to the effects of the immediate membrane environment during translocation (6), and the major receptor for this signal has been identified as the transmembrane potential (7, 8). In eukaryotes, published analyses of the effects of flanking charge on insertion orientation have focused principally on model type II membrane proteins. The use of type II proteins complicates interpretation because inverted Nexo insertion requires translocation of the prefolded native, normally cytoplasmic N-terminal domain. N-terminal domains preceding the first TM segment of integral membrane proteins can fold in the cytoplasm before translocation (6), and the major receptor for this signal has been identified as the transmembrane potential (7, 8).
flanking charge is much less, increasing the N-terminal charge. In the absence of the C-terminal RSRK motif, the orientation remained 94% Nexo, indicating a primary role for these negative charges in the topogenic signal. This was confirmed by additional N-terminal mutations in the truncated S42 series in which these negative charges were balanced by a positively charged terminus. The orientation receptor, a negatively charged terminus, promotes translocation of the most negatively charged terminus.

The charge difference and the positive inside rules predict the observed exclusive Nexo insertion of the S79-PB, S79a-PB, and S79-SPB fusions. The positive inside rule predicts cytoplasmic location for the RSRK motif C-terminal to TM1; its deletion resulted in partial inversion from Nocy to Nexo (13). The fractional N-terminal positive charge of this macrodipole, however, should favor Ncyt insertion.

Both the charge difference and the positive inside rules predict the observed exclusive Nexo insertion of the S79-PB, S79a-PB, and S79-SPB fusions. The positive inside rule predicts cytoplasmic location for the RSRK motif C-terminal to TM1; its deletion resulted in partial inversion from Nocy to Nexo (13). The fractional N-terminal positive charge of this macrodipole, however, should favor Ncyt insertion.

Our studies have both confirmed the dominant role of the charge difference in the topogenic signal determining TM protein insertion orientation and established the utility of this particular in vivo model. This now provides us with the ability to dissect these topogenic signals in more detail. For example, although we have not attempted to systematically determine the effect of distance between the TM segment and charge on signal strength, our results are consistent with the 15-residue window chosen by Hartmann et al. (13), although a distance-related gradient of effect can be anticipated. Identification of the receptor for this signal would clarify these issues (28). One can imagine that the signal recognition particle binds the hydrophobic core of nascent TM segments and presents this to the translocation machinery at the ER as a loop with its termini juxtaposed. The orientation receptor, a negatively charged patch on a component of the translocon adjacent to the pore, would respond to the sum of the charges adjacent to these termini, promoting translocation of the most negatively charged terminus.

REFERENCES

1. Dunlop, J., Jones, P. C., and Finbow, M. E. (1995) EMBO J. 14, 3609–3616
2. Gafvelin, G., and von Heijne, G. (1994) Cell 77, 401–412
3. Kutay, U., Ahnert-Hilger, G., Hartmann, E., Wiedemann, B., and Rapoport, T. A. (1995) EMBO J. 14, 217–223
4. von Heijne, G. (1994) Subcell. Biochem. 22, 1–19
5. von Heijne, G. (1992) J. Mol. Biol. 225, 487–494
6. Kristalik, L. I., and Cramer, W. A. (1995) FEBS Lett. 369, 140–143
7. Cao, G., Kuhn, A., and Dalbey, R. E. (1995) EMBO J. 14, 866–875
8. Andersson, H., and von Heijne, G. (1994) EMBO J. 13, 2267–2272
9. Sipos, L., and von Heijne, G. (1993) Eur. J. Biochem. 213, 1333–1340
10. Parks, G. D., and Lamb, R. A. (1991) Cell 64, 777–787
11. Parks, G. D., and Lamb, R. A. (1993) J. Biol. Chem. 268, 19101–19109
12. Beltzer, J. P., Fiedler, K., Fuhrer, C., Geffen, I., Handschin, C., Wessels, H. P., and Speiss, M. (1991) J. Biol. Chem. 266, 975–978
13. Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1999) Proc. Natl. Acad. Sci. U. S. A. 86, 5786–5790
14. Nakashima, H., and Nishikawa, K. (1992) FEBS Lett. 303, 141–146
15. Wallin, E., and von Heijne, G. (1995) Protein Eng. 8, 693–698
16. Denzer, A. J., Nabholz, C. E., and Spiess, M. (1993) EMBO J. 13, 6311–6317
17. Spiess, M. (1995) FEBS Lett. 369, 76–79
18. Sato, T., Sakaguchi, M., Mihara, K., and Omura, T. (1990) EMBO J. 9, 2381–2387
19. Cartwright, C. P., and Tipper, D. J. (1991) Mol. Cell. Biol. 11, 2620–2628
20. Komano, H., and Fuller, R. S. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 10752–10756
21. Cartwright, C. P., Li, Y., Zhu, Y. S., Kang, Y. S., and Tipper, D. J. (1994) Yeast 10, 497–508
22. Schandel, K. A., and Jenness, D. D. (1994) Mol. Cell. Biol. 14, 7245–7255
23. Cartwright, C. P., Zhu, Y. S., and Tipper, D. J. (1992) Yeast 8, 261–272
24. Redding, K., Holcomb, C., and Fuller, R. S. (1991) J. Cell. Biol. 113, 527–538
25. Nilsson, I., and von Heijne, G. (1993) J. Biol. Chem. 268, 5798–5801
26. Chaudhuri, B., and Stephan, C. (1992) FEBS Lett. 304, 41–45
27. Lankford, S. P., Cosson, P., Bonifacino, J. S., and Klausner, R. D. (1993) J. Biol. Chem. 268, 4814–4820
28. Dalbey, R. E., Kuhn, A., and von Heijne, G. (1995) Trends Cell Biol. 5, 380–383

Determinants of Membrane Protein Insertion Orientation

24633