Biophysical Journal
Article

Biophysical Society

Harmonizing Experimental Data with Modeling to Predict Membrane Protein Insertion in Yeast

Christopher J. Guerriero,[1,*] Yessica K. Gomez,[3] Grant J. Daskivich,[1] Karl-Richard Reutter,[1] Andrew A. Augustine,[1] Kurt F. Weiberth,[1] Kunio Nakatsukasa,[1,2] Michael Grabe,[3] and Jeffrey L. Brodsky[1]
[1]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania; [2]Division of Biological Science, Graduate School of Natural Sciences, Nagoya City University, Nagoya, Aichi, Japan; and [3]Cardiovascular Research Institute, Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California

ABSTRACT Membrane proteins must adopt their proper topologies within biological membranes, but achieving the correct topology is compromised by the presence of marginally hydrophobic transmembrane helices (TMHs). In this study, we report on a new model membrane protein in yeast that harbors two TMHs fused to an unstable nucleotide-binding domain. Because the second helix (TMH2) in this reporter has an unfavorable predicted free energy of insertion, we employed established methods to generate variants that alter TMH2 insertion free energy. We first found that altering TMH2 did not significantly affect the extent of protein degradation by the cellular quality control machinery. Next, we correlated predicted insertion free energies from a knowledge-based energy scale with the measured apparent free energies of TMH2 insertion. Although the predicted and apparent insertion energies showed a similar trend, the predicted free-energy changes spanned an unanticipated narrow range. By instead using a physics-based model, we obtained a broader range of free energies that agreed considerably better with the magnitude of the experimentally derived values. Nevertheless, some variants still inserted better in yeast than predicted from energy-based scales. Therefore, molecular dynamics simulations were performed and indicated that the corresponding mutations induced conformational changes within TMH2, which altered the number of stabilizing hydrogen bonds. Together, our results offer insight into the ability of the cellular quality control machinery to recognize conformationally distinct misfolded topomers, provide a model to assess TMH insertion in vivo, and indicate that TMH insertion energy scales may be limited depending on the specific protein and the mutation present.

SIGNIFICANCE Membrane proteins are difficult to fold because domain assembly must be coordinated between diverse environments, namely the endoplasmic reticulum lumen, cytoplasm, and lipid bilayer. To define how specific amino acids impact transmembrane helix (TMH) folding and insertion, we designed a dual-pass TMH reporter fused to a misfolded domain. Next, TMH integration and protein stability were measured in yeast. We observed a correlation between predicted insertion energy and TMH insertion, but some TMH variants outperformed predictions by knowledge-based or physics-based models. Based on molecular dynamics simulations, we propose that increased TMH insertion results from stabilizing hydrogen bonds. This finding highlights the need for further investigation of the properties that influence TMH insertion, with a focus on disease-causing mutations that alter TMH stability.

INTRODUCTION

Protein homeostasis relies on high fidelity synthesis and sorting of folded proteins to their functional sites. Proteins entering the secretory pathway in eukaryotes are translocated through the Sec61 translocon complex into the endoplasmic reticulum (ER) (1). However, several barriers impede protein folding, including genetic mutations, transcription or translational errors, defects in post-translational modifications, and failure to oligomerize, each of which is impacted further by cellular stress. Fortunately, these aberrant species are recognized by ER-associated molecular chaperones. They are then retrotranslocated from the ER, polyubiquitinated, and degraded by the 26S proteasome in the cytoplasm. This process is known as ER-associated degradation (ERAD) (2–7). The ERAD pathway also regulates the levels of select native proteins in response to metabolic cues (8).

The translocation and folding of multipass membrane proteins is particularly problematic. When the translocon

Submitted April 9, 2019, and accepted for publication July 8, 2019.
*Correspondence: cjg11@pitt.edu
Editor: Charles Deber.
https://doi.org/10.1016/j.bpj.2019.07.013

encounters a hydrophobic stretch of ~19–30 amino acids (9), these segments exit via a lateral gate and partition into the ER membrane (10). Using various models, the characteristics that drive transfer of a transmembrane helix (TMH) from translocon homologs into a lipid bilayer have been elucidated. Critical features include the positions of polar and charged residues in and adjacent to the TMH, TMH length, and overall hydrophobicity (11–17). For example, Ile, Leu, Val, Ala, Phe, and Met are favored in TMHs because of their lipophilicity, but polar and charged amino acids are increasingly disfavored as they near the center of a TMH (16,18). Curiously, a significant fraction of TMHs in multipass membrane proteins should fail to enter the bilayer based on predicted calculations of insertion free energy (16,19), but this hurdle is overcome by information transmitted from adjacent helices (16,19–24). Nevertheless, little is known about how TMH hydrophobicity affects the fate of a misfolded substrate.

Herein, we explored the impact of marginal TMH hydrophobicity on ERAD by using a model dual-pass protein fused to an unstable nucleotide-binding domain (NBD) that resides in a yeast ATP-binding cassette (ABC) transporter. TMH2 of the model protein has an unfavorable free energy for insertion, resulting in deposition of the NBD into the ER lumen. We first tested the hypothesis that altering TMH2 free energy will impact degradation of the model protein. Surprisingly, the topological arrangement of the model protein did not alter the degradation profile. Our model protein also presented us with the opportunity to examine TMH insertion in yeast. Using a series of TMH2 variants, we then discovered that particular variants have stronger than predicted effects on apparent insertion free energy. By combining physics-based energy calculations and molecular dynamics (MD) simulations, we attributed heightened membrane stability of these variants with altered helix conformation and increased hydrogen bonding.

MATERIALS AND METHODS

Yeast strains, plasmids, and plasmid construction

Yeast was maintained as described previously (25). Table S1 lists the *Saccharomyces cerevisiae* strains used in this study, and Table S2 lists the plasmids and primers. Chimera N* construction was described in detail elsewhere (26). To create Chimera N* variants, 40 bp primers were designed to include the desired codon change in TMH2 in the center of the primer. Mutagenesis was carried out by PCR overlap extension using the pCG28 template for pCG125, pCG126, pCG127, pCG130, pCG147, pCG148, and pCG149. For some double mutants, single mutants were used as a template; for pCG128, pCG126 served as a template; for pCG129, pCG127 served as a template; and for pCG131, pCG130 served as a template. Two fragments were then generated, a 5′ fragment (Fragment A) and 3′ fragment (Fragment B), each containing the mutation encoded in their respective primers. Fragments A and B were then used in a second PCR. The two fragments were added alone for five cycles to anneal, followed by the addition of primers at the extreme 5′ (pCG119) and 3′ (pCG128) ends of the annealed fragments to stich the fragments in another PCR cycle. The full length fragments were then digested and ligated into an empty 2μ vector downstream of the *PGK* promoter (empty version of pCG63) using the *Eco*RI and SacI restriction sites. The fragment A and B primer pairs are (5′ Fragment A/3′ Fragment A);(5′ Fragment B/3′ Fragment B), pCG131 (oCG119/168);(oCG169/128), pCG130 (oCG119/152);(oCG153/128), pCG129 (oCG119/166);(oCG167/128), pCG128 (oCG119/164);(oCG165/128), pCG127 (oCG119,150);(oCG151/128), pCG126 (oCG119/160);(oCG161/128), pCG125 (oCG119/162);(oCG163/128), pCG147 (oCG119/215);(oCG214/128), pCG148 (oCG119/213);(oCG212/128), and pCG149 (oCG119/217);(oCG216/128). The complete DNA sequence of all Chimera N* variants was confirmed using primers oKN54, oCG06, and oCG07.

Antibodies, immunoblot analysis, and indirect immunofluorescence microscopy

For immunoblot analysis, rat monoclonal anti-HA-horseradish peroxidase (HRP; 3F10; Roche, Basel, Switzerland) was used at 1:5000 to detect Chimera N*. As a loading control, rabbit anti-glucose-6-phosphate dehydrogenase (A9521; Sigma-Aldrich, St. Louis, MO) was used at 1:5000. For sucrose gradients, rabbit anti-Pma1p (Abcam, Cambridge, UK) was used at 1:2500, rabbit anti-Anp1p (a gift from Sean Munro, Cambridge University, Cambridge, United Kingdom) was used at 1:4000, and rabbit anti-Sec61p was used at 1:1000 (27). To detect myc-tagged ubiquitin, rabbit anti-myc (SC987; Santa Cruz Biotechnology, Dallas, TX) was used at 1:2500. After transfer to nitrocellulose, antibodies were incubated with blots overnight at 4°C, and bound primary antibodies were adorned with anti-rabbit immunoglobulin G HRP-conjugated secondary antibodies at 1:5000 for 2 h at room temperature. Bound antibodies were visualized using the SuperSignal Chemiluminescence kit (Thermo Fisher Scientific, Waltham, MA), images were captured using a Bio-Rad ChemiDoc XRS+ (Hercules, CA), and data were analyzed using ImageJ software, version 1.49b (National Institutes of Health). For indirect immunofluorescence microscopy, Chimera N* was detected with mouse anti-HA (12CA5; Roche) used at 1:500 and decorated with Alexa Fluor 488 goat anti-mouse at 1:500, rabbit anti-Kar2p (28) was used at 1:250 and detected with Alexa Fluor goat anti-rabbit at 1:500, and prolong antifade Gold with 4′,6-diamidino-2-phenylindole (Thermo Fisher) was used as a mounting agent and to detect nuclei. Indirect immunofluorescence was performed as described previously (29) and as reported recently (26) using an Olympus FV1000 (Olympus, Tokyo, Japan), ×100 UPlanSApo oil immersion objective, numerical aperture 1.40.

Cycloheximide chase assays

Protein turnover was measured using a cycloheximide (CHX) chase assay in yeast expressing Chimera N*. Cells were grown to OD_{600} = 0.5–1.5 in synthetic complete medium lacking uracil (−Ura) and containing 2% glucose. A 1 mL aliquot of cells was removed at the 0 min time point, and then CHX was added to a final concentration of 175 μg/mL. The culture was incubated in a shaking water bath at 26 or 37°C, as indicated, at 200 rpm. A 1 mL aliquot was removed at each time point into ice-cold tubes containing 0.5 M NaN_3 (final concentration 17.5 mM), and the cells were pelleted and flash frozen. Proteins were extracted by alkaline lysis and trichloroacetic acid (TCA) precipitation (30). The protein pellet was disrupted with a mechanical pestle in TCA sample buffer (80 mM Tris (pH 8), 8 mM EDTA, 3.5% SDS, 15% glycerol, 0.08% Tris base, 0.01% bromophenol blue) supplemented with fresh β-mercaptoethanol (final concentration 5%), and the samples were heated to 37°C for 30 min. An aliquot of each sample was subjected to sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) on a 10 or 12.5% polyacrylamide gel. Proteins were transferred onto nitrocellulose (BioTrace NT; Pall, Port Washington, NY) using a Trans-Blot Turbo Transfer System (Bio-Rad) and immunoblotted as described above.

Sucrose gradient analysis

Sucrose gradients were performed as described previously (26). Briefly, BY4742 cells expressing HA-tagged Chimera N* were grown to log phase, and ~40 OD_{600} equivalents of cells were pelleted in a clinical centrifuge and resuspended in 400 μL of 10 mM Tris (pH 7.6), 10 mM EDTA, 10% sucrose supplemented with 2 mM phenylmethlsulfonyl fluoride, 3 μg/mL leupeptin, 1.5 μg/mL pepstatin A, and 1 mM dithiothreitol. Cells were frozen dropwise into liquid N_2 and stored at −80°C. The frozen cells were then subjected to liquid N_2 lysis by grinding in a prechilled mortar and pestle for ~5 min, refreshing the liquid N_2 once each minute. After thawing the yeast powder, unbroken cells were removed by centrifugation for 2 min at 2000 rpm in a microcentrifuge at 4°C. The cleared lysate was then layered on an 11 mL 20–70% stepwise sucrose gradient, and 0.5% of the lysate was retained as the load fraction. Next, the gradients were centrifuged at 100,000 × *g* in a Beckman SW41 rotor for 18 h at 4°C. Fractions were collected from top to bottom, and any pelleted proteins at the bottom of the tube were solubilized in TCA sample buffer. An aliquot of each fraction was mixed with TCA sample buffer and analyzed by SDS-PAGE and immunoblot analysis as described above.

Measurements of substrate ubiquitination

pdr5Δ yeast cells were transformed with a Chimera N* expression plasmid (construct 10) or an empty vector and pKN31 for expression of a Cu^{2+}-inducible myc-tagged ubiquitin. Cells were grown to log phase at 26°C and treated with either dimethylsulfoxide or 20 μM MG132 for 1 h concomitant with the addition of 100 μM copper sulfate. Cells were harvested in a clinical centrifuge and washed once with ice-cold water, and pellets were stored at −80°C. Cells were lysed as described (26), and Chimera N* was immunoprecipitated using anti-HA-conjugated agarose beads (Roche). Chimera N* was liberated from the beads with TCA sample buffer supplemented with β-mercaptoethanol and analyzed by SDS-PAGE and immunoblotting. Before antibody exposure, the nitrocellulose membrane was incubated in boiling water for 1 h to further denature polyubiquitin epitopes.

Continuum energy calculations

An idealized α-helix for the wild-type TMH2 variant in Table 2 was constructed in Avogadro 1.2.0 (31), followed by minimization with ProteinPrep Wizard in Schrodinger Maestro (32). From this structure, we used Pymol to create the mutants, holding all other atoms fixed during the procedure. The high probability rotamer was always selected. Next, PARSE atomic partial charges were set using PDB2PQR (33), and we neutralized the charges on the N- and C-termini by hand. Briefly, our physics-based membrane insertion energy (ΔG) is given by

$$\Delta G = \Delta G_{elec} + \Delta G_{np}, \quad (1)$$

where the electrostatic energy (ΔG_{elec}) was calculated by solving the linear Poisson-Boltzmann equation and the nonpolar energy (ΔG_{np}) was assumed to be proportional to the protein's solvent-exposed surface area with a surface tension of a = 0.023 kcal/mol/Å^2 (Table 1). The reference state for both energy contributions is the protein free in solution, and the final state is the membrane-embedded structure. ABPSmem was used to compute both energies using calls to APBS for the electrostatics (34) and MSMS for the protein surface area calculations (35), with a salt concentration of 100 mM. The contribution to the nonpolar energy was linearly scaled from 0 to a for portions of the protein spanning the headgroup region. See (36) for more details.

The optimal membrane configuration of each construct was identified by computing ΔG for a wide range of orientations and then selecting the minimal value. Initially, the center of mass of each helix was positioned at the origin, with the long axis of the helix aligned to the membrane normal (z axis). The helix was rotated about its long axis by an angle ψ ranging from 0 to 360° in 10° increments, and for each value of ψ, it was then tilted off the z axis ϕ about the point $\vec{P}$ = (0, 0, 21 Å) at the N-terminus. ϕ ranged from 0 to 90° in 10° increments. The energy was computed using the linearized Poisson-Boltzmann equation, with one level of focusing to a resolution of 0.51 Å/grid for each orientation producing heat maps like those shown in Fig. 4. From each insertion energy map, we identified the lowest energy configuration and redid the calculation using the nonlinear equation with a finer, final resolution of 0.31 Å/grid. These calculations do not properly capture the energy of the uninserted state at the membrane interface, which is undoubtedly much lower energy than having the hydrophobic segments free in solution. Therefore, we shifted all ΔG values by +26.9 kcal/mol before plotting in Fig. 5 *A*.

The theoretical percent insertion based on Boltzmann statistics for a two-state system with experimental energy values ΔG_{app} (*black curve*, Fig. 5 *A*) is given by

$$P(\Delta G_{app}) = \frac{1}{1.3 + e^{\frac{\Delta G_{app}}{k_B T}}} + 0.15, \quad (2)$$

where under strongly destabilizing energy changes, 15% of the inserted band is still observed, whereas under strongly stabilizing conditions, 8% of the protein still fails to insert.

TABLE 1 Parameters for Physics-Based Model

Force Field	PARSE
Counterions	±1|*e*|, 0.1 M, 2.0 Å
Temperature	298.15 K
Grid dimensions	161 × 161 × 161
Coarse grid size	300 × 300 × 300 Å^3
Fine grid size	50 × 50 × 50 Å^3
Protein dielectric	2
Membrane dielectric	2
Headgroup dielectric	80
Solvent dielectric	80
Hydrophobic thickness	26 Å
Headgroup thickness	8 Å
Grid center	origin
Solution method	npbe
Boundary condition	zero
Charge model	spl2
Surface model	mol
Spline width	0.3 Å
Solvent probe radius	1.4 Å
Surface sphere density	10 Å^{-2}
Nonpolar surface tension	23 cal/mol/Å^2

Molecular dynamics simulations

All-atom MD simulations of the wild-type, G10L, S13L, and P15L helices were initiated using the same TMH2 models generated for the continuum calculations. Each helix was embedded in a phosphatidylcholine lipid bilayer containing 50 lipids per leaflet and solvated in 150 mM KCl using CHARMM-GUI (37), resulting in an average box size of 28000 atoms. These were then equilibrated in GROMACS 2018.3 (33) using the default equilibration scheme provided by CHARMM-GUI (38) for a total of 1.0 ns each. Production simulations were run for 500 ns each using the CHARMM36 force field (39) with a semi-isotropic pressure tensor and the Parrinello-Rahman barostat using a 5 ps^{-1} piston frequency. Temperature coupling using a Nose-Hoover extended ensemble was used with a reference temperature of 303.15 K and piston frequency of 1 ps^{-1}. The

TABLE 2 Free Energies and Percentage of TMH2 Insertion for Chimera N* variants

Construct #	ΔG_{pred} (kcal/mol)	ΔG_{app} (kcal/mol)	TMH2 Sequence	% Insertion
1	−2.37	−1.23	RSMAVMAL**L**AA**L**V**L**VMWLSLTSW	87.6
2	−1.60	−1.75	RSMAVMALGAA**L**V**L**VMWLSLTSW	90.8
3	−1.12	−1.73	RSMAVMAL**L**AASV**L**VMWLSLTSW	91.4
4	−0.53	−1.91	RSMAVMAL**L**AA**L**VPVMWLSLTSW	92.1
5	−0.17	−1.29	RSMAVMALGAASV**L**VMWLSLTSW	88.9
6	0.15	−0.42	RSMAVMALGAA**L**VPVMWLSLTSW	66.1
7	0.40	−1.01	RSMAVMAL**L**AASVPVMWLSLTSW	81.0
8	0.42	−0.20	RSMAVMALGAA**M**VPVMWLSLTSW	58.4
9	0.57	0.47	RSMAVMALGAA**A**VPVMWLSLTSW	30.8
10	1.06	1.07	RSMAVMALGAASVPVMWLSLTSW	16.5
11	1.76	0.93	RSMAVMALGAA**N**VPVMWLSLTSW	16.9

The ΔG_{pred} values were determined using ΔG Predictor (http://dgpred.cbr.su.se), and the ΔG_{app} values were calculated by the equation $\Delta G_{app} = -RT\ln K_{app}$, where $K_{app} = (f_{inserted}/f_{uninserted})$ and T = 293 K. The sequence for each TMH2 variant is listed, with alterations from the native sequence (construct 10, ΔG_{pred} = 1.07 kcal/mol) indicated in bold font and underlined. The percent insertion for each TMH2 variant was determined from n = 2–4 independent experiments performed with technical replicates.

SHAKE algorithm was used with a 2 fs time step. A nonbonded cutoff of 12 Å was used, and electrostatics were calculated using the particle mesh Ewald method, with snapshots saved every 500 ps.

The distinct kinking behavior present in several simulations was initially identified by visual inspection. We then used the VMD Bendix plugin (version 1.1) to quantify these values by calculating the average local kink angle θ (defined in Fig. 5 *E*, *top*) along the length of the helix over the entire simulation (40). Using a side length of 3.6, Bendix moves down the helix and calculates θ for every 3.6 residues and assigns this value to the top residue of the group while assigning linearly increasing values to the first three peripheral axis residues and linearly decreasing values to the last three, whose kink angles cannot be calculated. To quantify the helix tilt angle relative to the membrane, we computed the angle ϕ (defined in Fig. 5 *E*, *middle*) between the helix and the membrane normal using

$$\phi = \cos^{-1}\left(\frac{\vec{R}\cdot\hat{z}}{|\vec{R}|}\right),$$

where $\vec{R}$ is the vector from the COM of residue 2 to the COM of residue 24 and $\hat{z}$ is the membrane normal vector. Mean ϕ values in Fig. 4 are averaged over the last 250 ns of each simulation. Finally, the total number of hydrogen bonds for residues 2–24 was calculated over the last 250 ns using the MDAnalysis with default parameters (41).

RESULTS

To study the effect of marginal hydrophobicity on ERAD, we generated a dual-pass protein with an appended C-terminal misfolded domain ("degron"), which is sufficient to trigger proteasome-mediated degradation (Figs. 1 *A* and S1) (26,42). The degron was modeled on an ABC transporter in *S. cerevisiae*, Sterile 6 (Ste6p) (43,44). Ste6p possesses two sets of 6 TMHs, each followed by a cytoplasmic NBD, which links ATP hydrolysis to the transport of the **a**-type mating factor (Fig. 1 *A*). A 42-amino-acid truncation in the C-terminal NBD2 results in ERAD, and the substrate was termed Ste6p* (45). As a conserved member of the ABC transporter family, Ste6p is predicted to have 12 TMHs (46); however, most of the TMHs in Ste6p have an unfavorable ΔG for insertion (Table S3, http://dgpred.cbr.su.se) (16,17). Using a series of invertase fusions, Geller et al. validated proper insertion of the first six TMHs of Ste6p (47). Therefore, we simplified our analysis by linking only TMH1-2 from Ste6p to the 42-amino-acid-truncated NBD2 (NBD2*). This species was termed Chimera N* (26). Although the native TMH2 in this chimeric protein segment was unable to insert into the ER membrane, the topology could be corrected by replacing TMH2 with a 20-amino-acid poly-Ala/Leu sequence (36). When adopting its proper (inserted) topology with NBD2* in the cytoplasm, Chimera N* does not utilize its single available consensus site for the addition of an N-glycan moiety (NxS/T). Notably, the N-glycan consensus site is present in the ER lumenal loop between TMH1 and TMH2 in Chimera N*; however, its proximity to the lipid bilayer prevents access to the oligosaccharyl transferase complex (48). In contrast, failed insertion of TMH2 deposits NBD2* into the ER lumen, granting access to four N-glycosylation consensus sites that are present in NBD2* (Fig. 1 *A*). This attribute provides a quantitative method to monitor THM2 insertion into the ER.

The predicted ΔG for insertion of TMH2 (ΔG_{pred}) calculated using ΔG Predictor (http://dgpred.cbr.su.se/) is 1.06 kcal/mol (16,17). By examining the electrophoretic mobility of Chimera N* after expression at low copy number (pCG28, Table S1) in wild-type yeast (BY4742, Table S2), the protein largely adopted the uninserted orientation. This was evidenced by a slower than predicted migration pattern compared to the unglycosylated species (~43 kDa; Fig. 1 *B*, *left*). We next treated cell extracts with the bacterial enzyme endoglycosidase H (Endo H), which removes N-glycan moieties. Digestion with Endo H confirmed that the slower-migrating species had acquired N-linked glycans because of aberrant localization of NBD2* into the ER lumen (Fig. 1 *B*, *right*). Based on the observed shift in mobility, all four N-glycan acceptor sites in NBD2* were used.

As noted above, replacing TMH2 with a poly-Ala/Leu sequence corrects TMH2 topology, resulting in a dual-pass

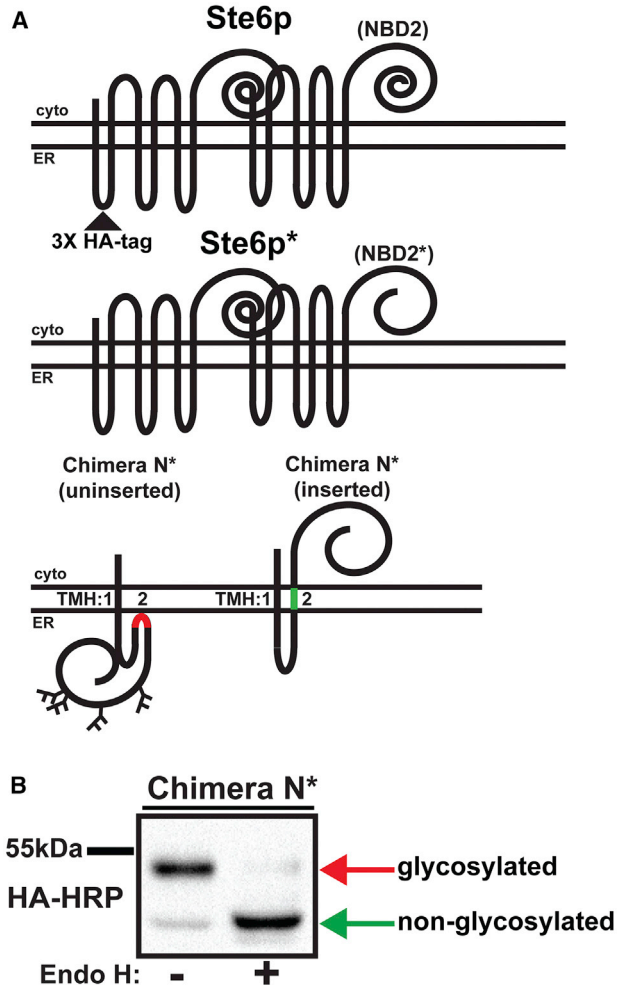


FIGURE 1 The design of a model substrate to measure transmembrane domain insertion in yeast. (*A*) A cartoon depicting the topology of Ste6p and Ste6p* and the two topologies adopted by Chimera N* is given. Each construct has a 3× HA-tag in the lumenal loop between TMH1 and TMH2. Chimera N* was made by internal deletion of amino acids 141–1042 in the coding sequence for 3× HA Ste6p*. Chimera N* displays dual topologies, with TMH2 residing inside the ER lumen (*left*, *red*), which consequently deposits NBD2* into the ER lumen and allows for N-glycan addition. Alternatively, TMH2 can insert properly into the membrane (*right*, *green*), which prevents acquisition of N-linked glycans. (*B*) *S. cerevisiae* strain BY4742 expressing Chimera N* under the control of a low-expression *ADH* promoter (pCG28) was grown to log phase, and cellular protein was extracted and incubated in the absence or presence of Endo H. Chimera N* was detected by SDS-PAGE and subsequent immunoblotting with anti-HA antibody. Red and green arrows to the right of the immunoblot mark the positions of the glycosylated and nonglycosylated forms, respectively.

substrate (with NBD2* in the cytosol) that is targeted for ERAD (26). Therefore, to test whether the single-pass substrate Chimera N* is also an ERAD substrate, we examined its fate. A hallmark of ERAD substrates is retention in the ER even though no peptide ER retention signals are present. This phenotype is typified by the disease-causing allele F508del in the cystic fibrosis transmembrane conductance regulator (CFTR), which is completely ER-retained (49).

We first used indirect immunofluorescence microscopy and confirmed Chimera N* residence in the ER based on a perinuclear and cortical staining pattern that is typical of the ER in yeast. In addition, Chimera N* staining overlapped with an ER-localized chaperone, Kar2p (Fig. 2 *A*). ER residence was further confirmed by sucrose gradient

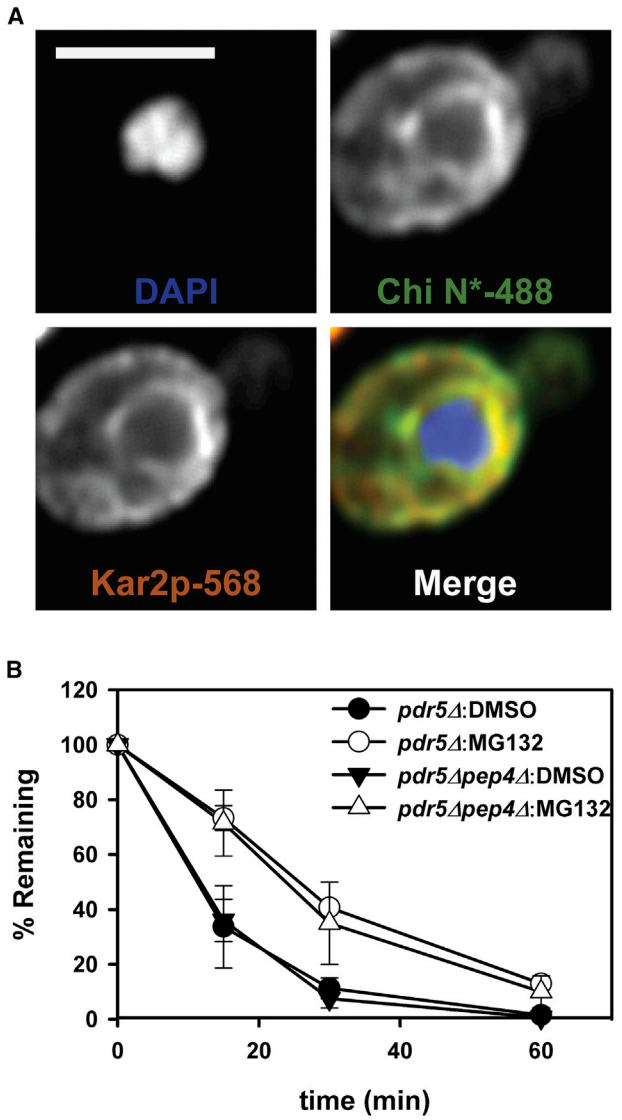


FIGURE 2 Chimera N* resides in the ER and is degraded by ERAD. (*A*) The residence of Chimera N* in the cell was investigated by indirect immunofluorescence using mouse anti-HA antibody (Chimera N*, *green*), rabbit anti-Kar2p (ER lumen, *red*), and 4′,6-diamidino-2-phenylindole (nuclei, *blue*). Primary antibodies were decorated with Alexa goat anti-mouse 488 and goat anti-rabbit 568. Images were captured using a confocal microscope, and a slice through the plane of the ER is shown. Scale bars, ~5μm. (*B*) Chimera N* was expressed at high copy in *pdr5Δ* (*circles*) or *pdr5Δpep4Δ* (*triangles*) yeast. Before CHX chase analysis, cells were preincubated at 26°C with either dimethylsulfoxide (*filled symbols*) or 100 μM MG132 (*open symbols*) for 20 min and chased for the indicated times. Graphed data represent the means ± SD from a representative experiment of n = 4 independent experiments for *pdr5Δ* and n = 1 performed with four technical replicates for *pdr5Δpep4Δ*.

analysis. Because the ER is bound by ribosomes engaged in translation, the ER can be isolated from the plasma membrane by differential centrifugation because of its increased density. We observed that Chimera N* comigrated with Sec61p but not with a plasma-membrane-resident protein (Fig. S2 *A*). To determine whether Chimera N* is an ERAD substrate, we next examined degradation via CHX chase. CHX is a bacterial toxin that binds the ribosome and halts new protein synthesis (50). Therefore, CHX treatment allows us to monitor the turnover of a pool of Chimera N* over time. The CHX chase was performed in a *pdr5Δ* yeast strain, which lacks a major drug pump and allows for treatment with a proteasome inhibitor, MG132 (51,52). As shown in Fig. 2 *B*, Chimera N* degradation was delayed in MG132-treated yeast, indicative of ERAD. Because Chimera N* was partially stabilized by treatment with MG132, we examined the potential contribution of yeast vacuolar proteases on Chimera N* degradation in a strain that also lacked vacuolar activity (*pdr5Δpep4Δ*). In some instances, a portion of a misfolded protein can be trafficked from the ER and degraded in the yeast vacuole, which is equivalent to the lysosome (53). However, we found that the overall degradation level was unchanged in *pdr5Δpep4Δ* yeast, suggesting no vacuolar involvement. Additionally, it has previously been observed that high-level expression of an ERAD substrate can redirect a portion of the misfolded protein to the vacuole (54). However, comparable results were obtained regardless of whether the substrate was expressed at high or low levels (Fig. S2, *B* and *C*).

A critical step in targeting a substrate to the proteasome is the addition of a polyubiquitin chain by E3 ubiquitin ligases (2). Therefore, we next examined Chimera N* polyubiquitination and observed an increase in the polyubiquitin ladder after treatment with MG132 (Fig. S2 *D*). Taken together, our data demonstrate that despite its unanticipated topology, Chimera N* is an ERAD substrate.

To explore the contributions of diverse TMHs on ER membrane insertion and ERAD, we designed Chimera N* variants using *ΔG* Predictor to determine ΔG_{pred}. The *ΔG* Predictor tool is a knowledge-based energy scale (herein referred to as a bioinformatics scale) built from in vitro studies on the insertion of artificial TMHs (16,17). To this end, we generated 10 Chimera N* variants with altered TMH2 sequences (Table 2). As noted above, the ΔG_{pred} for TMH2 in Chimera N* is 1.06 kcal/mol (Table 2, construct 10). We then focused on altering residues near the predicted middle of TMH2 because deleterious residues in this position should have the strongest effect on insertion (16). We next expressed the variants in yeast and quantified insertion (Fig. 3 *A*; Table 2). Although the Chimera N* degron (NBD2*) remains constant in each variant, as ΔG_{pred} for TMH2 decreases, the location of NBD2* shifts from primarily the ER lumen to the cytoplasm. Therefore, we anticipated different degradation rates for the two forms of Chimera N* because the ER lumen and cytosol contain unique chaperones and quality control factors. Unexpectedly, we found that the degradation profiles of the inserted and uninserted forms of construct 8 (ΔG_{pred} = 0.42 kcal/mol)—which is the most conformationally mercurial variant (58% insertion)—were indistinguishable (Fig. S3 *A*). The same result was evident when comparing the inserted forms of construct 1 and construct 8 (Fig. S3 *B*). However, a modestly slower degradation rate was observed when we compared the uninserted forms of construct 8 with construct 10, which contains the wild-type TMH2, or construct 11, which contains the least hydrophobic TMH2 (Fig. S3 *C*). This kinetic effect was only observed at the extreme end of the tested hydrophobicity range and may represent altered chaperone protein recognition of the unfavorable helix deposited into the ER lumen.

The use of Chimera N* to decipher effects of TMH hydrophobicity on topology is reminiscent of a previous model based on an *Escherichia coli* leader peptidase reporter, which possesses 2 TMHs and an engineered "H-segment." The use of this substrate helped define how amino acid composition and position impact TMH insertion (16,22,55–57). Many of these studies were performed in dog pancreas microsomes. Although cell-free protein translation systems provide valuable insights, the rates of protein translation are slower in vitro than in cells, and the composition of key cytoplasmic factors that affect protein targeting to the ER, translocation, and quality control differ. Thus, it

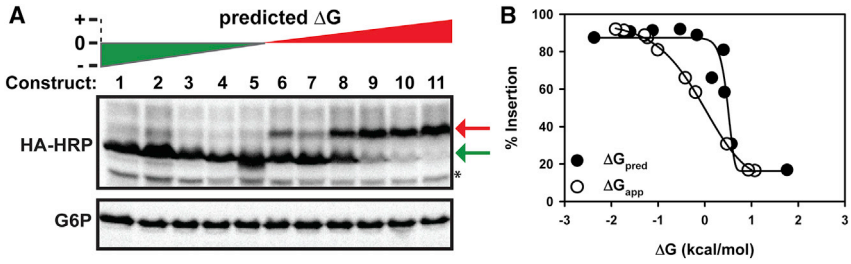


FIGURE 3 The predicted and measured insertion of Chimera N* in yeast. (*A*) *S. cerevisiae* expressing Chimera N* variants was grown to log phase, cellular proteins were extracted, and the Chimera N* species were resolved by SDS-PAGE and immunoblotted to detect the HA tag. Variants were loaded from left (*negative ΔG*, *green triangle*) to right (*positive ΔG*, *red triangle*) in the order listed in Table 2. The red and green arrows to the right mark the positions of the noninserted and inserted forms, respectively. * denotes a background band that migrates beneath the inserted form. A representative HA-HRP blot is shown, and the corresponding glucose-6-phosphate dehydrogenase blot serves as a loading control from n = 2–4 independent experiments. The percentage insertion was determined by the following equation: % inserted = $(f_{inserted}/(f_{inserted} + f_{uninserted})) \times 100$. (*B*) The data from part (*A*), displayed in Table 2, were plotted against ΔG_{pred} and ΔG_{app} and then fitted to a five parameter sigmoidal equation using SigmaPlot.

was vital to determine how ΔG_{pred} values correlate with the experimentally determined apparent free energy (ΔG_{app}) values for a substrate in its native environment, i.e., in yeast. We found that many TMH2 alterations predicted to have a minimal effect on ΔG_{pred} actually had a large impact on the ΔG_{app}, resulting in a much larger spread in ΔG_{app} values. For example, construct 10 (ΔG_{pred} = 1.06 kcal/mol) and construct 7 (ΔG_{pred} = 0.40 kcal/mol) insertion efficiencies were ~17% and 81%, respectively, although they differ only by the presence of a Gly or Leu at position 10. Moreover, when the percent insertion versus ΔG_{pred} and ΔG_{app} was examined using a nonlinear regression, the point at which 50% of TMH2 is inserted (I_{50}) corresponded to a ΔG_{pred} of 0.49 kcal/mol instead of the anticipated 0 kcal/mol, based on Boltzmann statistics used to compute the experimentally derived apparent free energy (Fig. 3 *B*). The discordance between the predicted and apparent ΔG values might reflect the significance of performing this study in vivo and highlights the importance of analyzing insertion parameters of TMHs in native environments.

To gain a better understanding of the insertion free energies of the TMH2 variants, we explored membrane stability with a physics-based model. In the absence of any structural information, we created idealized α-helices corresponding to each sequence in Table 2 and then computed their stabilities in the membrane compared to solution using a model that captures protein electrostatics and nonpolar stabilization in the membrane (26,36). We performed a scan over all angular orientations and identified the optimal inserted state (Fig. 4). These calculated energy values (Table S4) were all shifted by a constant offset to account for the uninserted state at the membrane interface and then plotted (*green dots*) along with ΔG_{pred} (*black circles* from Fig. 3 *B*) in Fig. 5 *A* (*bottom*). The physics-based calculations span nearly 10 kcal/mol in predicted stabilization energies, whereas ΔG_{pred} spans only 4 kcal/mol. Importantly, the former calculations more closely match the range of expected insertion probability values based on a two-state Boltzmann distribution (*black curve*), whereas the bioinformatics values jump from poorly inserted to highly inserted over a very narrow energy range. This discrepancy between scales is not surprising because bioinformatics scales often produce single amino acid insertion energies that are much smaller than physics-based values (58).

Nonetheless, two TMH2 variants are more stable in the membrane than predicted by either the bioinformatics or the physics-based models: G10L and P15L. To understand why these helices are so stable, we carried out all-atom MD simulations on the segments as well as the wild-type segment embedded in a phosphatidylcholine bilayer. It became immediately clear that the helices adopt different structural conformations. Most striking, the wild-type and G10L helices adopt a ~20° kink in the middle of the membrane at P15, in excellent agreement with previous simulation studies of proline-containing helices (59), but this kink is absent in P15L (*top panels*, Fig. 5, *B–E*, and *insets* of Fig. 5 *A*). Straightening the helix allows P15L to make 1.6 more backbone hydrogen bonds throughout the membrane spanning region than the wild-type TMH (*bottom panels*, Fig. 5, *B* and *C*), which is more than the expected value of 1 for Pro removal. Although G10L still kinks, it also makes almost 0.5 more hydrogen bonds on average than the wild-type segment, and this increase likely arises from removing the flexible Gly. Membrane hydrogen bonds have the potential to be very strong in the low dielectric of the membrane (60), and MD therefore supports our experimental observation that G10L and P15L stably integrate into the ER. The three segments also adopt different conformations in the membrane with G10L more aligned along the membrane normal (z axis), P15L adopting a 39° angle with respect to the z axis, and the wild-type segment exhibiting a more moderate angle (29°, *middle panels*, Fig. 5, *B–E*).

The physics-based model predicted the S13L mutation to be the most stabilizing of all single-point mutations, and the experiments revealed a greater than 40% increase in insertion probability over the wild-type. This large stabilization arises because the optimal membrane insertion places the polar serine in the middle of the membrane, and serine has a much higher insertion free energy than leucine, consistent with detailed free-energy calculations on isolated amino acid side chains (58) and biology-based scales (16,61). Nonetheless, we wanted to explore why S13L was less stabilizing than G10L or P15L. Thus, we also carried out

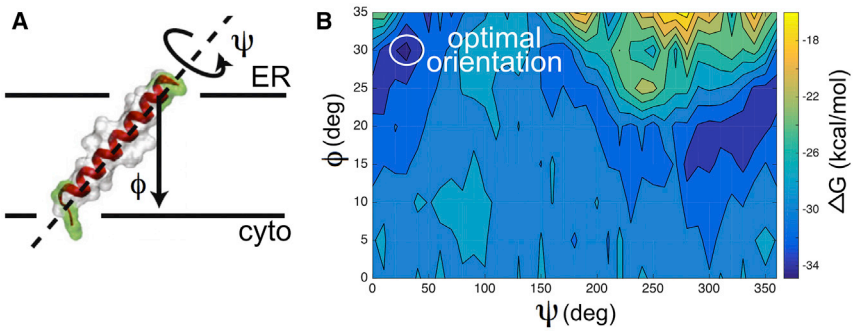


FIGURE 4 Calculation of optimal insertion orientation. (*A*) Definition of angles describing helix orientation in the membrane is given. (*B*) Insertion energy of the physics-based continuum model based on Eq. 1 of the supplement for the wild-type helix is shown as a function of helix rotation (ψ) and tilt angle (ϕ). The minimal energy value (*circle*) used to compute energies in Fig. 5 *A* is highlighted. The energy barrier corresponding to embedding the end of the helix in the membrane can be seen from the high energy values from ϕ = 45–70° (*yellow*).

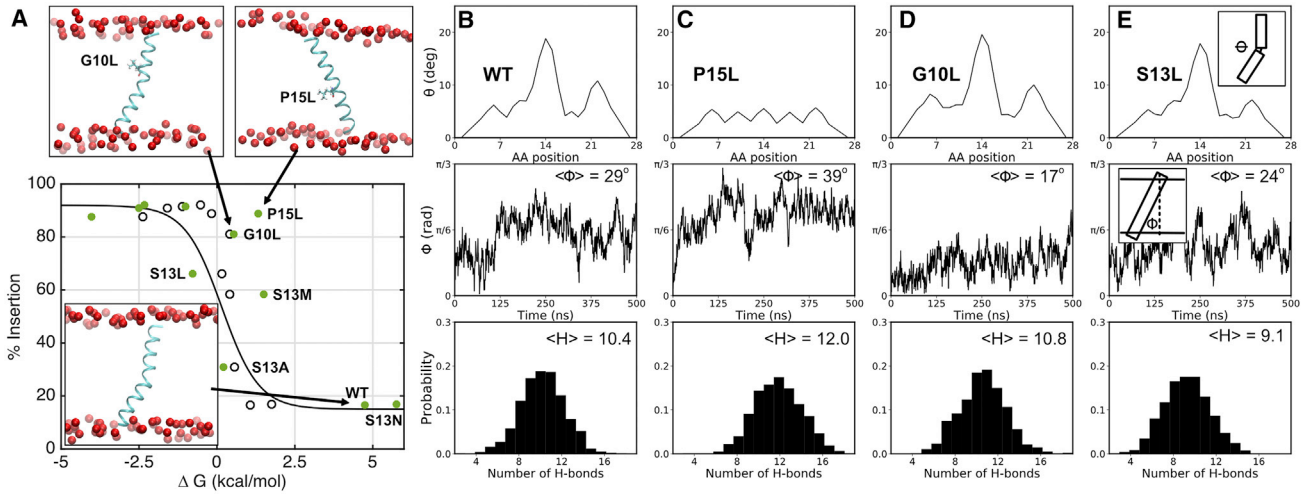


FIGURE 5 Physical models of insertion stability and molecular simulations. (*A*) Energy values were derived from our physics-based energy model computed for each TMH2 variant and plotted against the measured percent insertion (*green dots*). ΔG_{pred} data are from Fig. 3 *B* (*black circles*). Percent insertion predicted for a two-state Boltzmann distribution with energy difference ΔG bounded between 15 and 92% is shown (*black curve*). Insets show representative MD snapshots of the wild-type (WT) and P15L and G10L variants with mutated residues shown as licorice, protein backbone in a new cartoon, and headgroup phosphate atoms as red spheres. (*B–E*) Simulation details for the wild-type (*B*), P15L (*C*), G10L (*D*), and S13L (*E*) simulations are shown. Top: the average kink angle θ (defined in last panel) along the helix was computed, revealing large kinks at position 14 for wild-type and G10L and a straight helix for P15L. Middle: the angle ϕ that the helix makes with the membrane normal (defined in *last panel*) is shown. The average over the last half of each simulation is shown in the upper left corner. Bottom: a histogram of total backbone hydrogen bonds through the membrane span over the trajectory is shown. The average value computed over the last half of the simulation is shown in the upper left corner.

all-atom MD simulations on S13L to determine whether the removal of polar side chains in the core dramatically impacted the helix. Like the other proline-containing structures, S13L kinks at P15 (*top panel*, Fig. 5 *E*), and the orientation is similar to the wild-type (*middle panel*, Fig. 5 *E*). Not surprisingly, the loss of hydrogen-bonding capabilities of the serine side chain upon mutation to leucine reduced the average hydrogen bonding of the entire helix by a little more than one hydrogen bond compared to the wild-type, two compared to G10L, and three compared to P15L (*bottom panel*, Fig. 5, *B–E*). Therefore, we believe that the stabilization of S13L is largely due to better electrostatics and an increased nonpolar component, but the reduced hydrogen bonding offsets the stability compared to G10L and P15L. Overall, our simulations reveal that the mutations influence the conformation and membrane insertion in a manner that is not fully captured in either our physical model or in the bioinformatic scale. The physics-based model also fails to predict the stabilizing influence of S13M, and this may result from several sources, including overestimating the electrostatics of methionine in our continuum electrostatic calculations, undersampling of protein conformational changes, or details of the uninserted conformation.

DISCUSSION

Herein, we report on a model dual-pass membrane protein that acts both as a reporter for membrane protein quality control and TMH insertion. For the most topologically frustrated Chimera N* variant with ~50% TMH2 insertion, we observed equivalent ERAD efficiencies between the inserted and uninserted forms (Fig. S3 *A*). This result was surprising, given previous observations that topologically distinct isoforms of aquaporin 1 and rhodopsin are differentially engaged by the cellular proteostasis machinery (62,63). For example, TMH7 in rhodopsin is marginally hydrophobic because of a large number of polar and charged residues, and altering residues in the middle of TMH7 dramatically impacted the total and cell surface levels of rhodopsin, presumably through degradation of aberrant topomers (63). In contrast, Chimera N* is a quality control substrate regardless of whether TMH2 was inserted or uninserted. Therefore, any differences in degradation most likely stem from differential recognition of NBD2* by ER lumenal or cytoplasmic quality control machinery. An additional difference between our work and previous studies is that Chimera N* has only two TMHs as opposed to six (for aquaporin 1) or seven (for rhodopsin). The relatively low complexity of Chimera N* may allow the ERAD machinery to act more efficiently to remove both the uninserted and inserted forms. Another consideration is the route by which some integral membrane ERAD substrates are degraded. For Chimera N* (construct 10), the majority of the protein adopts a single-pass topology with a marginally hydrophobic TMH1 as the only membrane anchor (Table S3). Previously, it was demonstrated that single-pass proteins with marginal TMHs can completely pass into the ER lumen before they are retrotranslocated to the cytoplasm and degraded (64). However, this does not appear to be the case for Chimera N* because the protein remains

membrane-associated in the presence of sodium carbonate, and degradation remains robust in the absence of a functional copy of the ER lumenal Hsp70, BiP (unpublished data). Interestingly, for the Chimera N* variants in which TMH2 was primarily uninserted, we observed different degradation rates (Fig. S3 *C*). As noted above, in this case, decreased degradation as TMH2 becomes less favorably inserted may result from altered chaperone recognition.

Our study also represents the first analysis, to our knowledge, of TMH insertion for a dual-pass yeast transmembrane protein in its native environment. As anticipated, there is an overall correlation between the ΔG_{pred} and ΔG_{app} values (Pearson correlation coefficient, R = 0.80). However, the absolute magnitude of the apparent free-energy changes is much larger, which is in better agreement with simulation/physics-based scales (36,58). Additionally, we noted discordance when comparing ΔG_{pred} and ΔG_{app} values at 50% percent insertion (Fig. 3 *B*). These differences might reflect the contributions of endogenous cytoplasmic factors and/or a more rapid rate of translation in vivo, both of which could impact TMH insertion. In fact, a previous study comparing the ΔG_{app} in vitro and in BHK cells measured a difference between microsome and cell-based data of ~ -0.5 kcal/mol (17). We observed a remarkably similar shift in I_{50} between the ΔG_{pred} and ΔG_{app}, (0.49 kcal/mol; Fig. 3 *B*). Our data are also in good agreement with a previous study examining a leader peptidase reporter (described above) expressed in yeast. The addition of a single Leu to the engineered H-segment reduced ΔG_{app} by 1.8 kcal/mol (from 0.46 to −1.34 kcal/mol) (57), which is comparable to the 1.3 kcal/mol reduction (construct 10 versus 6) we observed.

Although ΔG Predictor provides a powerful tool to analyze insertion energies of TMHs, there are some limitations in its predictive power. For example, ΔG Predictor fails to identify Ste6p's TMH6 as an inserted helix because of the presence of a high number of polar and charged residues (Table S3). Thus, TMH6 insertion likely depends on interactions with neighboring helices, as shown for other ABC transporters (65–67). We note that many of the rules that govern membrane protein insertion have been defined using model proteins with mostly favorable TMHs (16,22,55–57), but these studies do not take into account complex interactions between stably integrated TMHs and those with positive ΔG_{pred} values. For example, although Ste6p has 12 TMHs, only TMH4 ($\Delta G_{pred} = -1.20$ kcal/mol) and 10 ($\Delta G_{pred} = -1.25$ kcal/mol) have favorable free energies of insertion (Table S3). Therefore, the insertion of distinct Ste6p TMHs likely depends on intimate interactions with neighboring helices to achieve the proper topology. In addition to interhelix interactions, it was also recently shown that the 100-amino-acid-residue region C-terminal to a marginally hydrophobic TMH influenced insertion efficiency, indicating that the conformation of a nascent polypeptide as it emerges from the ribosome can influence existing TMHs (68). Overall, the rules governing the insertion of TMHs into biological membranes continue to evolve.

Many disease-causing integral membrane proteins have been linked to ERAD (8), including aquaporin 2, polycystin 2, CFTR, and rhodopsin, all of which have mutant variants in predicted TMHs (69–72). For channel proteins, such as CFTR, TMH mutations can disrupt critical pore residues necessary for ion permeation (73,74), but they can also disrupt pore architecture, folding, and maturation (71,75). Therefore, our data have implications for understanding the factors that influence TMH insertion in both normal and diseased states. Furthermore, our findings highlight the need for more studies utilizing a combination of computational and in vivo approaches to examine the relationship between positional effects of amino acids and adjacent TMHs on insertion.

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2019.07.013.

AUTHOR CONTRIBUTIONS

C.J.G., J.L.B., K.N., and M.G. designed the research. C.J.G., G.J.D., Y.K.G., K.-R.R., A.A.A., and K.F.W. performed the experiments. C.J.G, J.L.B., and M.G. analyzed the data and wrote the manuscript.

ACKNOWLEDGMENTS

We thank the members of the Brodsky laboratory for their thoughtful insights on experimental design and feedback during the preparation of this manuscript, and especially Olivia Hilal for technical assistance, Jen Goeckeler-Fried and Markus Deserno for critical reading of the manuscript, and Thomas Kleyman for helpful advice and conversations.

This work was supported by National Institutes of Health grants DK101584 to C.J.G. and R25GM056847-19 to the UCSF Initiative for Maximizing Student Development Program supporting Y.K.G., a Beckman Scholars award to K.F.W., and National Institutes of Health grants GM117593 to M.G. and GM75061, DK079307, and R35 GM131732 to J.L.B.

SUPPORTING CITATIONS

References (76,77) appear in the Supporting Material.

REFERENCES

1. Cymer, F., G. von Heijne, and S. H. White. 2015. Mechanisms of integral membrane protein insertion and folding. *J. Mol. Biol.* 427:999–1022.
2. Vembar, S. S., and J. L. Brodsky. 2008. One step at a time: endoplasmic reticulum-associated degradation. *Nat. Rev. Mol. Cell Biol.* 9:944–957.
3. Hampton, R. Y., and T. Sommer. 2012. Finding the will and the way of ERAD substrate retrotranslocation. *Curr. Opin. Cell Biol.* 24:460–466.
4. Olzmann, J. A., R. R. Kopito, and J. C. Christianson. 2013. The mammalian endoplasmic reticulum-associated degradation system. *Cold Spring Harb. Perspect. Biol.* 5:a013185.

5. Christianson, J. C., and Y. Ye. 2014. Cleaning up in the endoplasmic reticulum: ubiquitin in charge. *Nat. Struct. Mol. Biol.* 21:325–335.
6. Pisoni, G. B., and M. Molinari. 2016. Five questions (with their answers) on ER-associated degradation. *Traffic.* 17:341–350.
7. Berner, N., K. R. Reutter, and D. H. Wolf. 2018. Protein quality control of the endoplasmic reticulum and ubiquitin-proteasome-triggered degradation of aberrant proteins: yeast pioneers the path. *Annu. Rev. Biochem.* 87:751–782.
8. Guerriero, C. J., and J. L. Brodsky. 2012. The delicate balance between secreted protein folding and endoplasmic reticulum-associated degradation in human physiology. *Physiol. Rev.* 92:537–576.
9. Baeza-Delgado, C., M. A. Marti-Renom, and I. Mingarro. 2013. Structure-based statistical analysis of transmembrane helices. *Eur. Biophys. J.* 42:199–207.
10. Rapoport, T. A., L. Li, and E. Park. 2017. Structural and mechanistic insights into protein translocation. *Annu. Rev. Cell Dev. Biol.* 33:369–390.
11. Spiess, M., T. Junne, and M. Janoschke. 2019. Membrane protein integration and topogenesis at the ER. *Protein J.* 38:306–316.
12. Junne, T., L. Kocik, and M. Spiess. 2010. The hydrophobic core of the Sec61 translocon defines the hydrophobicity threshold for membrane integration. *Mol. Biol. Cell.* 21:1662–1670.
13. White, S. H., and W. C. Wimley. 1994. Peptides in lipid bilayers: structural and thermodynamic basis for partitioning and folding. *Curr. Opin. Struct. Biol.* 4:79–86.
14. Wimley, W. C., and S. H. White. 1996. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.* 3:842–848.
15. Liu, L. P., S. C. Li, …, C. M. Deber. 1996. Threshold hydrophobicity dictates helical conformations of peptides in membrane environments. *Biopolymers.* 39:465–470.
16. Hessa, T., N. M. Meindl-Beinker, …, G. von Heijne. 2007. Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature.* 450:1026–1030.
17. Hessa, T., H. Kim, …, G. von Heijne. 2005. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature.* 433:377–381.
18. Zhu, C., Y. Gao, …, X. C. Zeng. 2016. Characterizing hydrophobicity of amino acid side chains in a protein environment via measuring contact angle of a water nanodroplet on planar peptide network. *Proc. Natl. Acad. Sci. USA.* 113:12946–12951.
19. Ojemalm, K., K. K. Halling, …, G. von Heijne. 2012. Orientational preferences of neighboring helices can drive ER insertion of a marginally hydrophobic transmembrane helix. *Mol. Cell.* 45:529–540.
20. Hedin, L. E., K. Ojemalm, …, A. Elofsson. 2010. Membrane insertion of marginally hydrophobic transmembrane helices depends on sequence context. *J. Mol. Biol.* 396:221–229.
21. Buck, T. M., J. Wagner, …, W. R. Skach. 2007. A novel tripartite motif involved in aquaporin topogenesis, monomer folding and tetramerization. *Nat. Struct. Mol. Biol.* 14:762–769.
22. Meindl-Beinker, N. M., C. Lundin, …, G. von Heijne. 2006. Asn- and Asp-mediated interactions between transmembrane helices during translocon-mediated membrane protein assembly. *EMBO Rep.* 7:1111–1116.
23. Zhang, L., Y. Sato, …, N. Uozumi. 2007. Contribution of hydrophobic and electrostatic interactions to the membrane integration of the Shaker K+ channel voltage sensor domain. *Proc. Natl. Acad. Sci. USA.* 104:8263–8268.
24. Lerch-Bader, M., C. Lundin, …, G. von Heijne. 2008. Contribution of positively charged flanking residues to the insertion of transmembrane helices into the endoplasmic reticulum. *Proc. Natl. Acad. Sci. USA.* 105:4127–4132.
25. Adams, A., D. E. Gottschling, …, T. Stearns. 1997. Methods in Yeast Genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
26. Guerriero, C. J., K. R. Reutter, …, J. L. Brodsky. 2017. Transmembrane helix hydrophobicity is an energetic barrier during the retrotranslocation of integral membrane ERAD substrates. *Mol. Biol. Cell.* 28:2076–2090.
27. Stirling, C. J., J. Rothblatt, …, R. Schekman. 1992. Protein translocation mutants defective in the insertion of integral membrane proteins into the endoplasmic reticulum. *Mol. Biol. Cell.* 3:129–142.
28. Brodsky, J. L., and R. Schekman. 1993. A Sec63p-BiP complex from yeast is required for protein translocation in a reconstituted proteoliposome. *J. Cell Biol.* 123:1355–1363.
29. Amberg, D. C., D. Burke, and J. N. Strathern. 2005. Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
30. Nakatsukasa, K., G. Huyer, …, J. L. Brodsky. 2008. Dissecting the ER-associated degradation of a misfolded polytopic membrane protein. *Cell.* 132:101–112.
31. Hanwell, M. D., D. E. Curtis, …, G. R. Hutchison. 2012. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* 4:17.
32. Schrödinger LLC. 2016. Small-Molecule Drug Discovery Suite 2016–1: Maestro, version 10.5; LigPrep, version 3.7; Epik, version 3.5; Glide, version 7.0; Primer, version 4.3.
33. Dolinsky, T. J., J. E. Nielsen, …, N. A. Baker. 2004. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 32:W665–W667.
34. Baker, N. A., D. Sept, …, J. A. McCammon. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA.* 98:10037–10041.
35. Sanner, M. F., A. J. Olson, and J. C. Spehner. 1996. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers.* 38:305–320.
36. Marcoline, F. V., N. Bethel, …, M. Grabe. 2015. Membrane protein properties revealed through data-rich electrostatics calculations. *Structure.* 23:1526–1537.
37. Jo, S., T. Kim, …, W. Im. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29:1859–1865.
38. Lee, J., X. Cheng, …, W. Im. 2016. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* 12:405–413.
39. Huang, J., and A. D. MacKerell, Jr. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* 34:2135–2145.
40. Dahl, A. C., M. Chavent, and M. S. Sansom. 2012. Bendix: intuitive helix geometry analysis and abstraction. *Bioinformatics.* 28:2193–2194.
41. Michaud-Agrawal, N., E. J. Denning, …, O. Beckstein. 2011. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* 32:2319–2327.
42. Guerriero, C. J., K. F. Weiberth, and J. L. Brodsky. 2013. Hsp70 targets a cytoplasmic quality control substrate to the San1p ubiquitin ligase. *J. Biol. Chem.* 288:18506–18520.
43. Kuchler, K., R. E. Sterne, and J. Thorner. 1989. Saccharomyces cerevisiae STE6 gene product: a novel pathway for protein export in eukaryotic cells. *EMBO J.* 8:3973–3984.
44. McGrath, J. P., and A. Varshavsky. 1989. The yeast STE6 gene encodes a homologue of the mammalian multidrug resistance P-glycoprotein. *Nature.* 340:400–404.
45. Loayza, D., A. Tam, …, S. Michaelis. 1998. Ste6p mutants defective in exit from the endoplasmic reticulum (ER) reveal aspects of an ER quality control pathway in Saccharomyces cerevisiae. *Mol. Biol. Cell.* 9:2767–2784.
46. Decottignies, A., and A. Goffeau. 1997. Complete inventory of the yeast ABC proteins. *Nat. Genet.* 15:137–145.

47. Geller, D., D. Taglicht, …, E. Bibi. 1996. Comparative topology studies in Saccharomyces cerevisiae and in Escherichia coli. The N-terminal half of the yeast ABC protein Ste6. *J. Biol. Chem.* 271:13746–13753.

48. Nilsson, I. M., and G. von Heijne. 1993. Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. Biol. Chem.* 268:5798–5801.

49. Cheng, S. H., R. J. Gregory, …, A. E. Smith. 1990. Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell.* 63:827–834.

50. Schneider-Poetsch, T., J. Ju, …, J. O. Liu. 2010. Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat. Chem. Biol.* 6:209–217.

51. Lee, D. H., and A. L. Goldberg. 1996. Selective inhibitors of the proteasome-dependent and vacuolar pathways of protein degradation in Saccharomyces cerevisiae. *J. Biol. Chem.* 271:27280–27284.

52. Gaczynska, M., and P. A. Osmulski. 2005. Small-molecule inhibitors of proteasome activity. *Methods Mol. Biol.* 301:3–22.

53. Sun, Z., and J. L. Brodsky. 2018. The degradation pathway of a model misfolded protein is determined by aggregation propensity. *Mol. Biol. Cell.* 29:1422–1434.

54. Spear, E. D., and D. T. Ng. 2003. Stress tolerance of misfolded carboxypeptidase Y requires maintenance of protein trafficking and degradative pathways. *Mol. Biol. Cell.* 14:2756–2767.

55. Whitley, P., I. M. Nilsson, and G. von Heijne. 1996. A nascent secretory protein may traverse the ribosome/endoplasmic reticulum translocase complex as an extended chain. *J. Biol. Chem.* 271:6241–6244.

56. Sääf, A., E. Wallin, and G. von Heijne. 1998. Stop-transfer function of pseudo-random amino acid segments during translocation across prokaryotic and eukaryotic membranes. *Eur. J. Biochem.* 251:821–829.

57. Lundin, C., H. Kim, …, G. von Heijne. 2008. Molecular code for protein insertion in the endoplasmic reticulum membrane is similar for N(in)-C(out) and N(out)-C(in) transmembrane helices. *Proc. Natl. Acad. Sci. USA.* 105:15702–15707.

58. MacCallum, J. L., W. F. Bennett, and D. P. Tieleman. 2007. Partitioning of amino acid side chains into lipid bilayers: results from computer simulations and comparison to experiment. *J. Gen. Physiol.* 129:371–377.

59. Cordes, F. S., J. N. Bright, and M. S. Sansom. 2002. Proline-induced distortions of transmembrane helices. *J. Mol. Biol.* 323:951–960.

60. Cao, Z., J. M. Hutchison, …, J. U. Bowie. 2017. Backbone hydrogen bond strengths can vary widely in transmembrane helices. *J. Am. Chem. Soc.* 139:10742–10749.

61. Moon, C. P., and K. G. Fleming. 2011. Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proc. Natl. Acad. Sci. USA.* 108:10174–10177.

62. Buck, T. M., and W. R. Skach. 2005. Differential stability of biogenesis intermediates reveals a common pathway for aquaporin-1 topological maturation. *J. Biol. Chem.* 280:261–269.

63. Roushar, F. J., T. C. Gruenhagen, …, J. P. Schlebach. 2019. Contribution of cotranslational folding defects to membrane protein homeostasis. *J. Am. Chem. Soc.* 141:204–215.

64. Feige, M. J., and L. M. Hendershot. 2013. Quality control of integral membrane proteins by assembly-dependent membrane integration. *Mol. Cell.* 51:297–309.

65. Enquist, K., M. Fransson, …, I. Nilsson. 2009. Membrane-integration characteristics of two ABC transporters, CFTR and P-glycoprotein. *J. Mol. Biol.* 387:1153–1164.

66. Carveth, K., T. Buck, …, W. R. Skach. 2002. Cooperativity and flexibility of cystic fibrosis transmembrane conductance regulator transmembrane segments participate in membrane localization of a charged residue. *J. Biol. Chem.* 277:39507–39514.

67. Tector, M., and F. U. Hartl. 1999. An unstable transmembrane segment in the cystic fibrosis transmembrane conductance regulator. *EMBO J.* 18:6290–6298.

68. Junne, T., and M. Spiess. 2017. Integration of transmembrane domains is regulated by their downstream sequences. *J. Cell Sci.* 130:372–381.

69. Bichet, D. G., A. El Tarazi, …, P. Bissonnette. 2012. Aquaporin-2: new mutations responsible for autosomal-recessive nephrogenic diabetes insipidus-update and epidemiology. *Clin. Kidney J.* 5:195–202.

70. Watnick, T., N. He, …, Y. Pei. 2000. Mutations of PKD1 in ADPKD2 cysts suggest a pathogenic effect of trans-heterozygous mutations. *Nat. Genet.* 25:143–144.

71. Patrick, A. E., A. L. Karamyshev, …, P. J. Thomas. 2011. Alteration of CFTR transmembrane span integration by disease-causing mutations. *Mol. Biol. Cell.* 22:4461–4471.

72. Illing, M. E., R. S. Rajan, …, R. R. Kopito. 2002. A rhodopsin mutant linked to autosomal dominant retinitis pigmentosa is prone to aggregate and interacts with the ubiquitin proteasome system. *J. Biol. Chem.* 277:34150–34160.

73. Sheppard, D. N., D. P. Rich, …, M. J. Welsh. 1993. Mutations in CFTR associated with mild-disease-form Cl- channels with altered pore properties. *Nature.* 362:160–164.

74. Hämmerle, M. M., A. A. Aleksandrov, and J. R. Riordan. 2001. Disease-associated mutations in the extracytoplasmic loops of cystic fibrosis transmembrane conductance regulator do not impede biosynthetic processing but impair chloride channel stability. *J. Biol. Chem.* 276:14848–14854.

75. Cui, G., Z. R. Zhang, …, N. A. McCarty. 2008. Mutations at arginine 352 alter the pore architecture of CFTR. *J. Membr. Biol.* 222:91–106.

76. Winzeler, E. A., D. D. Shoemaker, …, R. W. Davis. 1999. Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science.* 285:901–906.

77. Tsirigos, K. D., C. Peters, …, A. Elofsson. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res.* 43:W401–W407.

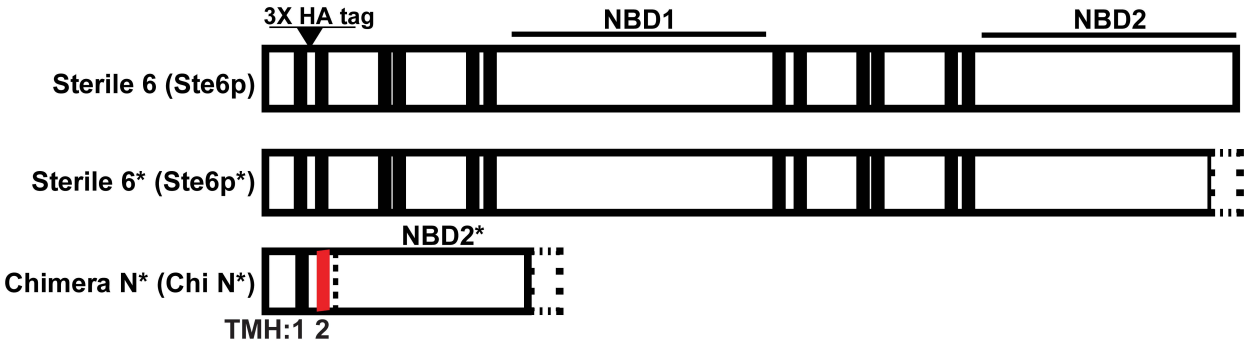
Biophysical Journal, Volume *117*

Supplemental Information

Harmonizing Experimental Data with Modeling to Predict Membrane Protein Insertion in Yeast

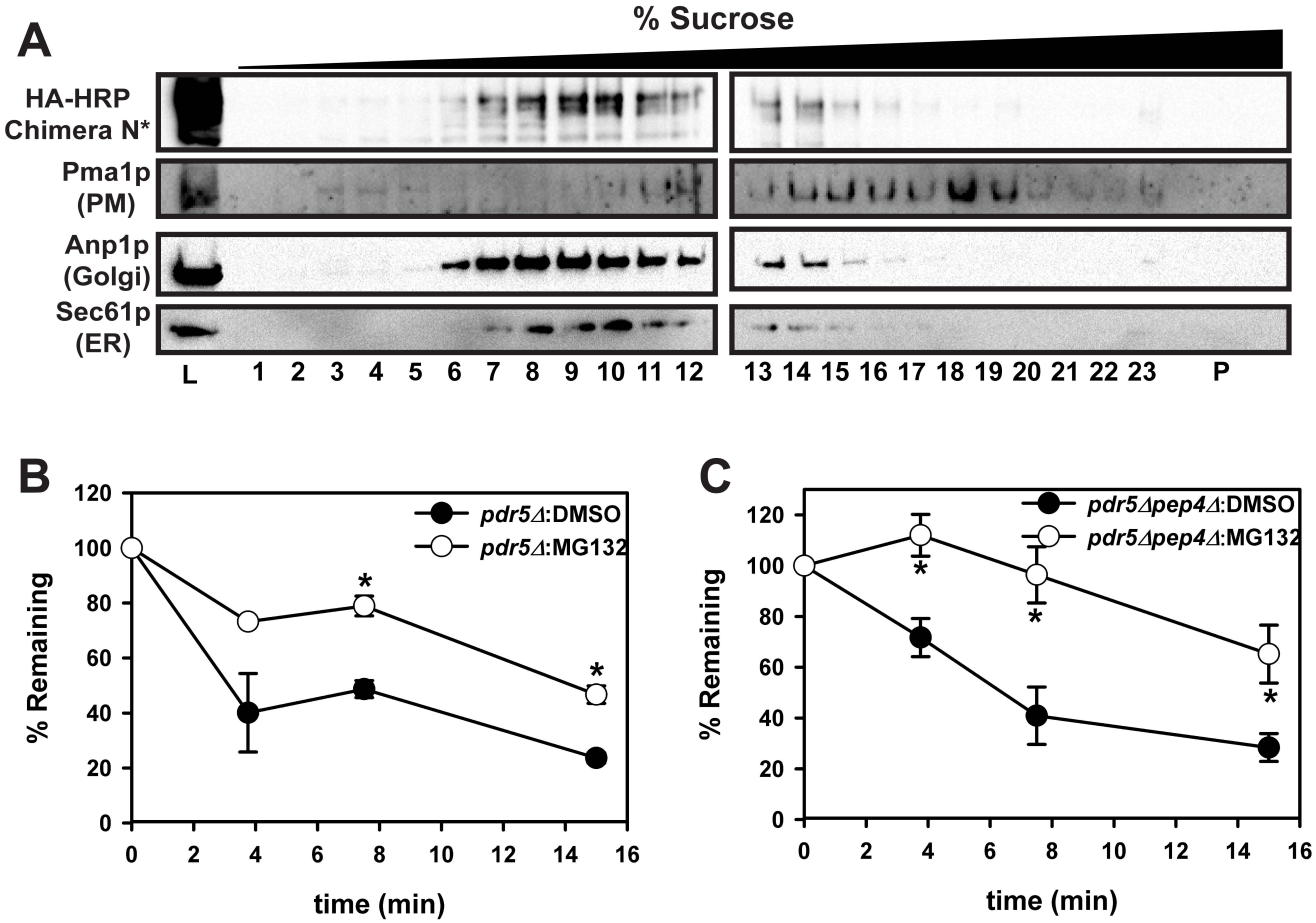
Christopher J. Guerriero, Yessica K. Gomez, Grant J. Daskivich, Karl-Richard Reutter, Andrew A. Augustine, Kurt F. Weiberth, Kunio Nakatsukasa, Michael Grabe, and Jeffrey L. Brodsky

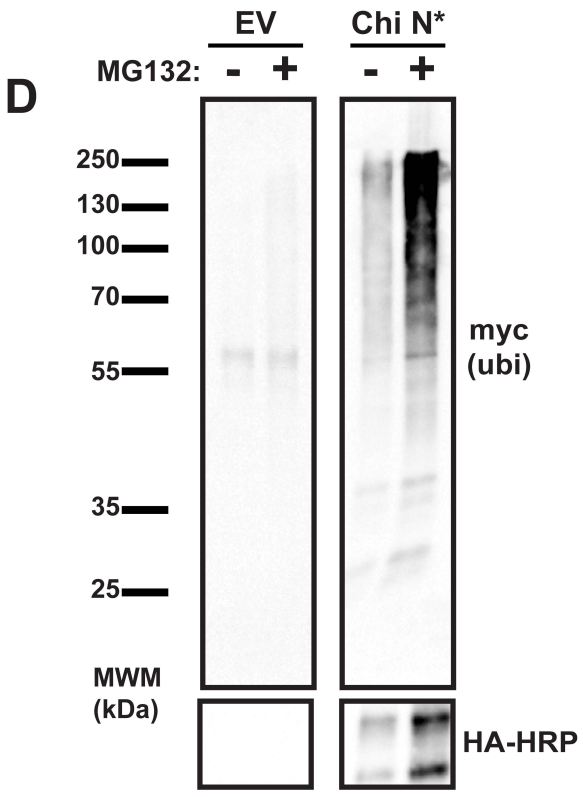
Supplemental Figure 1
Guerriero et al.



Supplemental Figure 1 Chimera N* is based on Ste6p*, an established ERAD substrate
Linear diagrams of Sterile 6 (Ste6p), the ERAD substrate Ste6p*, and Chimera N*. Ste6p is a yeast ABC transporter and contains 12 TMHs (black bars) and two nucleotide-binding domains (NBDs). The truncations in Ste6p* and Chimera N* are depicted as empty rectangles with a dashed border. The fusion point between the N-terminal and C-terminal portions of Ste6p* is depicted by a vertical dashed line immediately following TMH2 (colored red due its inability to insert into the membrane).

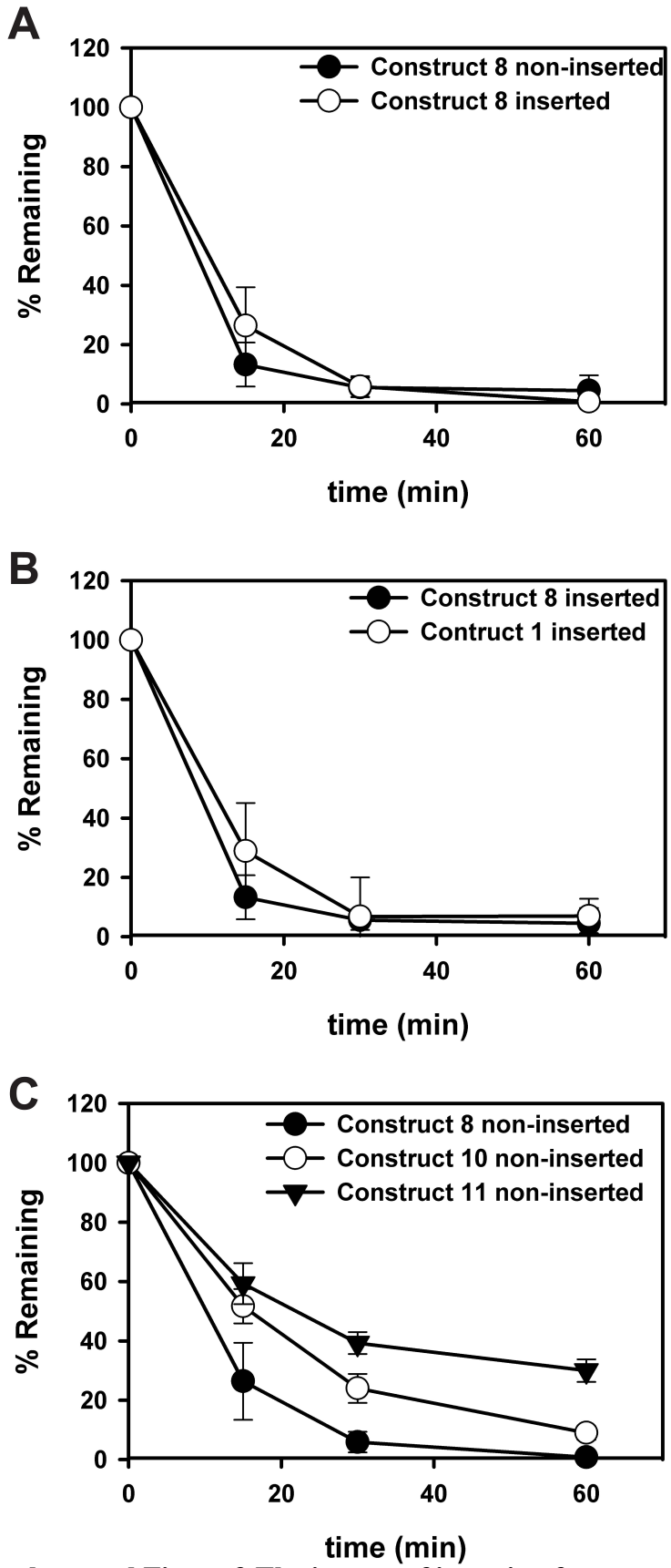
Supplemental Figure 2
Guerriero et al.





Supplemental Figure 2 Chimera N* resides in the ER and is degraded by the proteasome

A) Lysates from yeast expressing Chimera N* were subjected to sucrose gradient centrifugation and fractions were collected from the top (low % sucrose) to the bottom (high % sucrose). An aliquot from each fraction was analyzed by SDS-PAGE and immunoblotting for Chimera N* (HA-HRP), as well as Pma1p (a plasma membrane protein), Anp1p (a Golgi resident), and Sec61p (the ER translocation channel). A lane containing 0.5% of the total lysate is included at the left (L) and the pelleted material at the bottom of the tube was loaded at the far right (P). Chimera N* was expressed under the control of the ADH promoter (pCG28) in **(B)** *pdr5Δ* or **(C)** *pdr5Δpep4Δ* for analysis by cycloheximide chase. Cells were preincubated for 20 min with DMSO (closed circles) or 100 μM MG132 (open circles) and then chased for the indicated times. *p<0.03. **D)** *pdr5Δ* cells expressing Chimera N* and Cu2+-inducible myc-tagged ubiquitin were treated for 90 min with DMSO (-) or 50 μM MG132 (+). Chimera N* was immunoprecipitated with anti-HA conjugated agarose beads, followed by SDS-PAGE and immunoblotting for the myc tag and HA tag.



Supplemental Figure 3 The impact of insertion free energy on Chimera N* degradation kinetics
Chimera N* variants 1, 8, 10, and 11 were expressed under the control of the PGK promoter in wild type yeast (BY4742). Cells were treated with cycloheximide at the beginning of the chase and aliquots were removed at the indicated time points. Graphed data represent the means +/- SD from a representative experiment of n=2 independent experiments, each with four technical replicates. The degradation rates of the uninserted (glycosylated form) and inserted (unglycosylated form) are compared in A, B, and C, as indicated.

Supplemental Table 1 Primers and Plasmids used in this study

Primer/Plasmid	Sequence/Genotype	Reference
Primer		
oKN54	GATCATCAAGGAAGT	(26)
oCG06	AGGCAAAATTAAAATAGA	(26)
oCG07	TCTATTTTAATTTTGCCT	(26)
oCG119	GGAATTCCTGCAGCCCGGGATGAACTTTTTAAGTTTTAAG	This study
oCG128	ACCGAGCTCCACCGCGGTTATTCACTATGCGTTATAACCA	This study
oCG150	TTGTTAGAGAAAGCCACATTACCAACACAGAAGCCGCACC	This study
oCG151	CAATGGCAGTAATGGCACTTGGTGCGGCTTCTGTGTTGGT	This study
oCG152	TTGTTAGAGAAAGCCACATTACCAACACCAAAGCCGCACC	This study
oCG153	CAATGGCAGTAATGGCACTTGGTGCGGCTTTGGTGTTGGT	This study
oCG160	ACTGGCACAAGAGCCGCACCAAGTGCCATTACTGCCATTG	This study
oCG161	CAATGGCAGTAATGGCACTTGGTGCGGCT**CTT**GTGCCAGT	This study
oCG162	ACTGGCACAGAAGCCGCAAGAAGTGCCATTACTGCCATTG	This study
oCG163	CAATGGCAGTAATGGCACTTCTTGCGGCTTCTGTGCCAGT	This study
oCG164	ACTGGCACAGAAGCCGCAAGAAGTGCCATTACTGCCATTG	This study
oCG165	CAATGGCAGTAATGGCACTT**CT**TGCGGCT**CTT**GTGCCAGT	This study
oCG166	ACCAACACAGAAGCCGCAAGAAGTGCCATTACTGCCATTG	This study

oCG167	CAATGGCAGTAATGGCACTT**CT**TGCGGCTTCTGTG**TTG**GT	This study
oCG168	ACCAACACCAAAGCCGCCAAAAGTGCCATTACTGCCATTG	This study
oCG169	CAATGGCAGTAATGGCACTT**TTG**GCGGCT**TTG**GTG**TTG**GT	This study
oCG212	CAATGGCAGTAATGGCACTTGGTGCGGCT**GCT**GTGCCAGT	This study
oCG213	ACTGGCAC**AGC**AGCCGCACCAAGTGCCATTACTGCCATTG	This study
oCG214	CAATGGCAGTAATGGCACTTGGTGCGGCT**ATG**GTGCCAGT	This study
oCG215	ACTGGCAC**CAT**AGCCGCACCAAGTGCCATTACTGCCATTG	This study
oCG216	CAATGGCAGTAATGGCACTTGGTGCGGCT**AAT**GTGCCAGT	This study
oCG217	ACTGGCACATTAGCCGCACCAAGTGCCATTACTGCCATTG	This study
Plasmid		
pKN31	2µ *HIS3 Pcup1-mycUb-Tcyc1*	(30)
pCG28	CEN ADH *P-PGK ste6-166-3HA*	This study
pKN05/pCG63 (cons. 10)	2µ *URA3 ste6-166tm1-2*(ΔG=1.06)*NBD2-3HA*	(26)
pCG125 (cons. 7)	2µ URA3 *STE6*tm1-2(ΔG=0.40)NBD2-3HA	This study
pCG126 (cons. 6)	2µ URA3 *STE6*tm1-2(ΔG=0.15)NBD2-3HA	This study
pCG127 (cons. 5)	2µ URA3 *STE6*tm1-2(ΔG=-0.17)NBD2-3HA	This study
pCG128 (cons. 4)	2µ URA3 *STE6*tm1-2(ΔG=-0.53)NBD2-3HA	This study
pCG129 (cons. 3)	2µ URA3 *STE6*tm1-2(ΔG=-1.12)NBD2-3HA	This study
pCG130 (cons. 2)	2µ URA3 *STE6*tm1-2(ΔG=-1.60)NBD2-3HA	This study
pCG131 (cons. 1)	2µ URA3 *STE6*tm1-2(ΔG=-2.37)NBD2-3HA	This study
pCG147 (cons. 8)	2µ URA3 *STE6*tm1-2(ΔG=0.42)NBD2-3HA	This study

pCG148 (cons. 9)	2µ URA3 *STE6*tm1-2(ΔG=0.57)NBD2-3HA	This study
pCG149 (cons. 11)	2µ URA3 *STE6*tm1-2(ΔG=1.76)NBD2-3HA	This study

Supplemental Table 2 Yeast strains used in this study

Strain	Genotype	Reference/Source
BY4742	*MATα, his3Δ1, leu2Δ0, lys2Δ0 ura3Δ3*	(76)
pdr5Δ	*MATα, his3Δ1, leu2Δ0, ura3Δ3 pdr5Δ::KanMX*	(76)
pdr5Δ pep4Δ	*MATα, met15Δ0, his3Δ1, leu2Δ0, ura3Δ3 pdr5Δ::KanMX, pep4Δ::KanMX*	(26)

TMH	ΔG_{pred} (kcal/mol)	TMH2 sequence	Source
1	0.063	**YRLLMIMIIGTVATGLVPAIT**	dgpred
2	1.06	**RSMAVMALGAASVPVMWLSLTSW**	dgpred
3	2.28	**ASAITFQNLVAICALLGTS**	dgpred
4	-1.20	**FYYSWSLTLIILCSSPIITFFAV**	dgpred
5	0.041	**TFFIKSCFFVAANAGILRFLTLT**	dgpred
6	4.50	**NDVITCFHSCIMLGSTLNNTL**	TOPCONS
7	0.044	**IRYKKILILGLLCSLIAGAT**	dgpred
8	0.91	**YLAKWSLLVLGVAAADGIF**	dgpred
9	0.052	**FLSAMTSFVTVSTIGLIWALV**	dgpred
10	-1.25	**LSLVCISMFPLIIIFSAIYGGIL**	dgpred
11	0.92	**FGISMTNMIVMCIQAIIYYYGL**	dgpred
12	-0.071	**MFTTFTLLLFTIMSCTSLV**	dgpred

Supplemental Table 3 Predicted TMHs and free energy of insertion for Ste6p

Free insertion energies (ΔG kcal/mol) were determined by entering the amino acid sequence for 3XHA-Ste6p into ΔG Predictor – full protein scan. Since ΔG Predictor was unable to detect TMH6, an alternate search tool, TOPCONS (topcons.cbr.su.se) was utilized (77). Subsequently, ΔG Predictor was used to calculate the ΔG for the putative TMH6 sequence which was identified using TOPCONS.

Construct #	ΔG_{pred} (kcal/mol)	TMH2 sequence	% Insertion
1	-4.04	RSMAVMAL**L**AA**L**V**L**VMWLSLTSW	87.6
2	-2.49	RSMAVMALGAA**L**V**L**VMWLSLTSW	90.8
3	-1.02	RSMAVMAL**L**AASV**L**VMWLSLTSW	91.4
4	-2.32	RSMAVMAL**L**AA**L**VPVMWLSLTSW	92.1
5	1.32	RSMAVMALGAASV**L**VMWLSLTSW	88.9
6	-0.78	RSMAVMALGAA**L**VPVMWLSLTSW	66.1
7	0.54	RSMAVMAL**L**AASVPVMWLSLTSW	81.0
8	1.50	RSMAVMALGAA**M**VPVMWLSLTSW	58.4
9	0.21	RSMAVMALGAA**A**VPVMWLSLTSW	30.8
10	4.75	RSMAVMALGAASVPVMWLSLTSW	16.5
11	5.76	RSMAVMALGAA**N**VPVMWLSLTSW	16.9

Supplemental Table 4 Free energies values derived from continuum calculations

Free insertion energies (ΔG kcal/mol) were determined as described in Figure 4. All values have been shifted by a constant value of 26.9 kcal/mol to account for the uninserted state, for which we do not have structural information.