

Insertion of short transmembrane helices by the Sec61 translocon

Simon Jaud^a, Mónica Fernández-Vidal^{b,1}, IngMarie Nilsson^c, Nadja M. Meindl-Beinker^c, Nadja C. Hübner^c, Douglas J. Tobias^a, Gunnar von Heijne^c, and Stephen H. White^{b,2}

^aDepartment of Chemistry and the Center for Biomembrane Systems, University of California, Irvine, CA 92697-2025; ^bDepartment of Physiology and Biophysics and the Center for Biomembrane Systems, University of California, Irvine, CA 92697-4560; and ^cDepartment of Biochemistry and Biophysics and the Center for Biomembrane Research, Stockholm University, SE-106 91 Stockholm, Sweden

Edited by Donald M. Engelman, Yale University, New Haven, CT, and approved May 28, 2009 (received for review January 19, 2009)

The insertion efficiency of transmembrane (TM) helices by the Sec61 translocon depends on helix amino acid composition, the positions of the amino acids within the helix, and helix length. We have used an in vitro expression system to examine systematically the insertion efficiency of short poly-leucine segments (L_n , $n = 4 \dots 12$) flanked at either end by 4-residue sequences of the form $XXPX-L_n-XPXX$ with $X = G, N, D$, or K . Except for $X = K$, insertion efficiency (p) is $<10\%$ for $n < 8$, but rises steeply to 100% for $n = 12$. For $X = K$, p is already close to 100% for $n = 10$. A similar pattern is observed for synthetic peptides incorporated into oriented phospholipid bilayer arrays, consistent with the idea that recognition of TM segments by the translocon critically involves physical partitioning of nascent peptide chains into the lipid bilayer. Molecular dynamics simulations suggest that insertion efficiency is determined primarily by the energetic cost of distorting the bilayer in the vicinity of the TM helix. Very short lysine-flanked leucine segments can reduce the energetic cost by extensive hydrogen bonding with water and lipid phosphate groups (snorkeling) and by partial unfolding.

hydrophobic mismatch | membrane protein synthesis | membrane proteins | molecular dynamics simulation | lipid bilayer

The thickness of the hydrocarbon core of bilayer membranes is ≈ 30 Å, which leads to the expectation that transmembrane (TM) helices of helix-bundle membrane proteins (MPs) should be ≈ 20 -aa long to span the hydrocarbon core. Consistent with this expectation, the average helix length of α -helical MPs of known 3D structure is 26 ± 5 residues (<http://blanco.biomol.uci.edu/mptopo>). A few membrane proteins, most notably the aquaporins (1) and H^+/Cl^- exchangers (2) have membrane-buried helices as short as $n = 10$ residues, corresponding to a helix length of $(n - 1) \times 1.5 = 13.5$ Å. Although there is ample evidence that Sec61/SecY translocons can insert short helices, the biophysical basis of short-helix insertion has not been explored systematically.

Kuroiwa et al. (3) found that the Sec61 translocon could efficiently insert artificial stop-transfer poly-leucine sequences as short as 9 residues across rough microsomal membranes. Chen and Kendall (4) subsequently found that the *Escherichia coli* SecY translocon is also able to insert 10-residue poly-leucine segments across the inner membrane. In a detailed examination of the molecular code that the Sec61 translocon uses for identifying TM segments, Hessa et al. (5) found that efficient stop-transfer segments comprised of short poly-leucine helices could also be engineered into a multispinning membrane protein, specifically the *E. coli* leader peptidase (Lep). In a recent model-peptide study, Krishnakumar and London (6) determined the minimum hydrophobic length required for the formation of poly(Leu) or poly(LeuAla) TM helices in small unilamellar vesicles formed from diacyl phosphatidylcholines with chain acyl lengths of 14 to 24 carbons. Their results were consistent with the TM insertion of hydrophobic sequences as short as 11–12 leucine residues in vesicle membranes formed from dioleoylphosphatidylcholine (DOPC).

A limitation of the biochemical and model-peptide studies to date is that they can only be inferentially compared with each other.

To overcome this limitation, we have used the same set of TM helices (H-segments) for biochemical, biophysical, and computational studies. By using a microsomal, in vitro expression system to examine systematically the Sec61 insertion efficiency of short poly-leucine segments flanked at either end by 4-residue sequences of the form $XXPX-L_n-XPXX$ ($n = 4-12$; $X = G, N, D$, or K), we found that helices with $n = 10-12$ were readily inserted into the membrane. We also synthesized selected sequences from the $XXPX-L_n-XPXX$ family and examined by oriented circular dichroism (OCD) the incorporation of the peptides into oriented palmitoyl-oleoyl phosphocholine (POPC) and POPC:POPG (palmitoyl-oleoyl phosphatidylglycerol) multibilayers at various concentrations and hydrations. We found that short helices of the $XXPX$ family were indeed stable in a transmembrane configuration in oriented multibilayers. Finally, to explore the physical principles underlying the stability of short helices under conditions of severe hydrophobic mismatch, we carried out molecular dynamics (MD) simulations of some of the peptides in transmembrane orientations. Taken together, the results from this multifaceted approach suggest that the biophysics of hydrophobic mismatch controls the insertion of short helices across the endoplasmic reticulum (ER) membrane.

Results

Integration of Poly-leucine H-Segments Across ER Membranes by the Sec61 Translocon. We first determined insertion efficiencies of systematically designed H-segments of the form $XXPX-L_n-XPXX$ ($X = G, N, D$, or K) across the endoplasmic membrane by using the approach of Hessa et al. (5, 7), summarized in Fig. 1A. In brief, the H-segments were introduced near the middle of the large luminal P2 domain of the model protein, Lep. The protein was then expressed in vitro in the presence of ER-derived dog pancreas rough microsomes (RMs). The insertion probability (efficiency) p was calculated on the basis of the fractions of singly (f_{1g}) and doubly (f_{2g}) glycosylated proteins by using $p = f_{1g}/(f_{1g} + f_{2g})$ determined from quantitative phosphoimager scans of SDS/PAGE gels. Hessa et al. (5, 7) showed that p is proportional to the Boltzmann function $[1 + \exp(\Delta G_{app}/RT)]^{-1}$, which we consequently used to characterize the probability-of-insertion data.

The central L_n stretch was flanked by the tetrapeptides $XXPX$ - and $-XPXX$ to “insulate” it from the surrounding sequence in the Lep model protein. For the present work, we took H-segments with

Author contributions: I.N., D.J.T., G.v.H., and S.H.W. designed research; S.J., M.F.-V., I.N., N.M.M.-B., and N.C.H. performed research; S.J., M.F.-V., I.N., D.J.T., G.v.H., and S.H.W. analyzed data; and S.J., M.F.-V., D.J.T., G.v.H., and S.H.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹Present address: Department of Peptide and Protein Chemistry, Institut d'Investigacions Químiques i Ambientals de Barcelona, Consell Superior d'Investigacions Científiques, 08034 Barcelona, Spain.

²To whom correspondence should be sent at: Department of Physiology and Biophysics, University of California at Irvine, Medical Sciences I-D346, Irvine, CA 92697-4560. E-mail: stephen.white@uci.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/090638106/DCSupplemental.

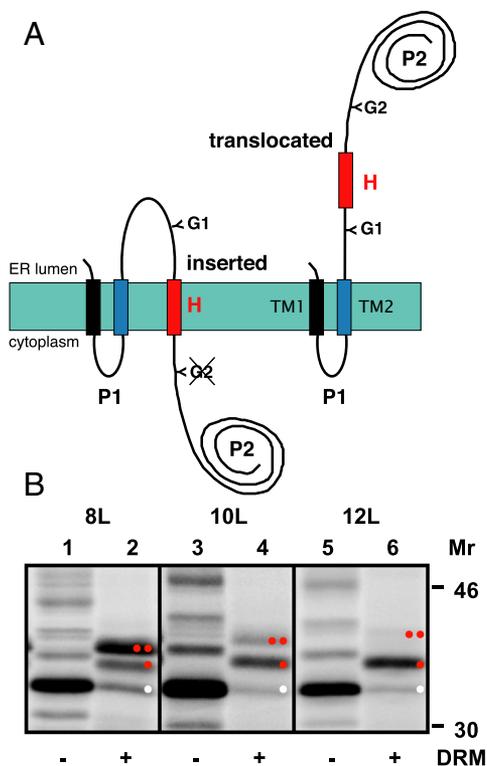


Fig. 1. Integration of polyleucine H-segments into microsomal membranes. (A) Leader peptidase (Lep) model protein construct. Wild-type Lep, consisting of 2 TM helices (TM1, TM2) and a large luminal domain (P2), inserts into rough microsomes in an N_{lum} - C_{lum} orientation. H-segments (red) of the form $XXPX-L_n$ - $XPXX$ with $n = 4$ to $n = 12$ and $X = G, N, D, \text{ or } K$ were engineered into the P2 domain between residues 226 and 253. Glycosylation acceptor sites (G1 and G2) were placed in positions 96–98 and 258–260, flanking the H-segment. For H-segments that integrate into the membrane, only the G1 site is glycosylated (*Left*), whereas both the G1 and G2 sites are glycosylated for H-segments that do not integrate into the membrane (*Right*). (B) Membrane integration of H-segments for GGPG- L_n -GPGG constructs with $n = 8, 10,$ and 12 . Plasmids encoding the Lep/H-segment constructs were transcribed and translated *in vitro* in the presence (+) and absence (-) of dog pancreas rough microsomes (DRM). Translation products were analyzed by SDS/PAGE. Bands of unglycosylated protein are indicated by white dots; singly (1g) and doubly (2g) glycosylated proteins are indicated by 1 and 2 red dots, respectively. The data were quantitated by scanning the gels in a phosphorimager. The probability (efficiency) of membrane insertion is given by $P = 1g/(1g + 2g)$. Mean values from 2 independent experiments were used for computing P values. On average, glycosylation levels vary by about $\pm 2\%$ between repeat experiments.

$X = G$ flanks as the reference segments. The immediate sequence environment outside the H-segment is quite polar (ENGIRLSETS/H-segment/VPGQONATWI), hence there is no additional hydrophobic flanking segment that could be “recruited” into the membrane-spanning H-segment. Examples of SDS/PAGE gels showing the translation products of GGPG- L_n -GPGG constructs with $n = 8, 10,$ and 12 are shown in Fig. 1B. Control experiments show that the identities of the TM1 and TM2 helices of the Lep constructs have little influence on the insertion efficiencies of the H-segment (8).

Fig. 2A shows plots of the probability of insertion as a function of n for $X = G, N, D,$ and K . The probabilities of insertion for the reference set of H-segments with $X = G$ are plotted as solid black squares. These data accurately conform to a Boltzmann distribution, shown as a solid black curve. This distribution means that the insertion process has the appearance of equilibrium between the inserted and translocated states measured by the relative amounts of singly and doubly glycosylated translation products. The data show that the probability of insertion is ≈ 0 for $n = 8$ or fewer

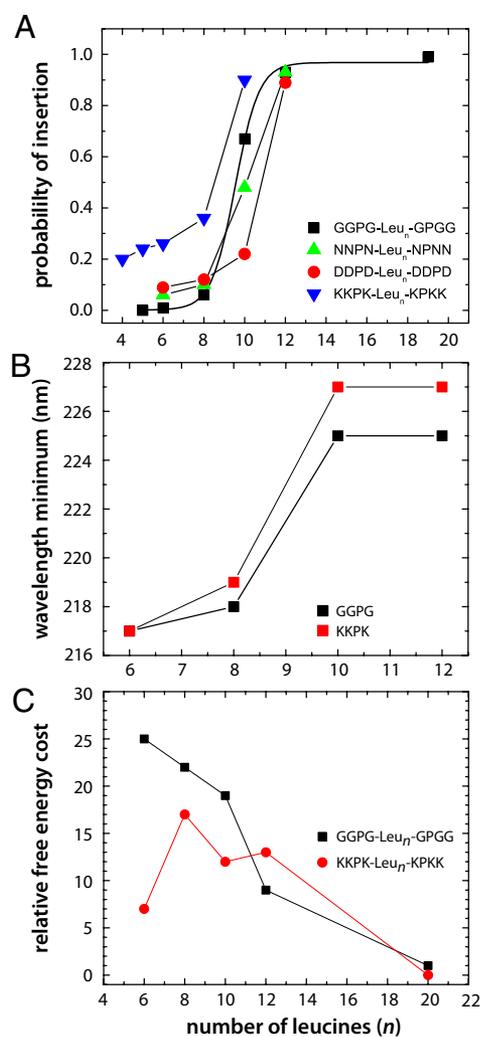


Fig. 2. Length dependence of various parameters. (A) Efficiency of insertion of $XXPX-L_n$ - $XPXX$ constructs in dog pancreas rough microsomes. Data for $X = G$ (black squares) were collected for $n = 5, 6, 8, 10, 12,$ and 19 . The solid curve through the points is a best-fit Boltzmann distribution (see *Results*). Diameters of points correspond to the experimental uncertainty. (B) Wavelength-minima of the OCD spectra as a function of the number of leucines for $X = G$ and $X = K$. Approximate representative positions of the minima are shown in Fig. 3B and D. Although the interpretation of the OCD spectra for $n \approx 8$ (Fig. 3) is uncertain, this plot reveals a structural change that approximates the shape of the probability-of-insertion curves shown here. All OCD spectra for $n \geq 10$ are consistent with TM α -helices, as shown by the examples in Fig. 3. (C) Relative free-energy costs of deforming a POPC bilayer to accommodate $XXPX-L_n$ - $XPXX$ peptides of different lengths and compositions. Relative free energy is the computed free energy for a given n divided by the free energy computed for GL20. By using the fits of Eq. S1 to the $d(r)$ curves of Fig. 5, we calculated the differences in deformation free energies of the lipid-protein system by using Eq. S2. We calculated the difference in free energy of all systems relative to data from a simulation of a GL20 peptide in a bilayer, $\Delta G_{peptide}/\Delta G_{GL20}$. GL20 was chosen as a reference because the hydrophobic mismatch is small. For this reason, the relative free energies for $n = 20$ are zero in this figure.

leucines, but rises steeply to $\approx 90\%$ for $n = 12$ or greater. These data confirm the earlier observations of Kuroiwa et al. (3) and Chen and Kendall (4). Assuming an α -helix, 12 leucines corresponds to a length of 16.5 Å, which cannot span the typical 30-Å-thick hydrocarbon core of lipid bilayers without considerable negative hydrophobic mismatch.

Replacement of the flanking Gly residues with Asn (Fig. 2, green triangles) had little effect on the probability of insertion, whereas

replacement with Asp (Fig. 2, red circles) caused a small shift of the insertion probabilities toward longer runs of leucines. In contrast, replacement of the flanking glycines with Lys residues (Fig. 2, inverted blue triangles) caused a shift of the probability curve to shorter lengths so that 90% efficiency was achieved for $n = 10$ leucines. If these results are due to physical interactions of the peptides with the membrane, then it should be possible to construct model membrane systems in which short synthetic peptides are stable across lipid bilayer membranes.

Integration of Synthetic Poly-leucine H-Segments Across Oriented Phospholipid Membranes. We synthesized 2 families of H-segments, each containing $n = 6, 8, 10,$ or 12 leucines. The glycine family (designated GL_n) had $X = \text{Gly}$ flanks; the lysine family (designated KL_n) had $X = \text{Lys}$ flanks. In methanol and trifluoroethanol solutions, the GL_n and KL_n peptide families were α -helical for all values of n as determined by solution CD spectra (Fig. S1). The conformations of the GL_n family in phosphate buffer depended on n (Fig. S2A). For $n = 6$ or 8 , the CD spectra were consistent with β -sheet structure, probably due to aggregation, whereas for $n = 10$ or 12 , the conformations were predominantly α -helical. In contrast, the CD spectra of the KL_n family in phosphate buffer were consistent with random coil for $n = 6$ or 8 and predominately α -helical for $n = 10$ or 12 (Fig. S2B). We conclude from these measurements that a conformational transition to α -helical occurs for both peptide families between $n = 8$ and $n = 10$.

The ability of the peptides to insert across a lipid bilayer was examined by means of OCD spectroscopy of peptides in oriented multibilayer arrays that were formed by depositing methanol solutions of lipid and peptide on quartz substrates. Highly oriented arrays form spontaneously after slow evaporation of the methanol followed by hydration via vapor-phase equilibration with saturated salt solutions at 66% or 93% relative humidity (RH) (9, 10). Although the bilayers are not fully hydrated at these RHs, the disposition of incorporated peptides is generally the same as in excess water (9, 11). Bilayers were composed of either POPC or 1:1 POPC:POPG; peptide:lipid concentrations ranged from 1:12 to 1:100. Bilayer/peptide multilayers (consisting of stacks of $\approx 1,000$ bilayers) formed in this way are highly aligned parallel to the quartz substrate as determined by X-ray diffraction.

OCD spectra were obtained for all peptides in both POPC and POPC:POPG at all peptide:lipid ratios (Fig. S3). Representative OCD data for the GL_n and KL_n peptide families ($n = 6$ or 12) are shown in Fig. 3. Shown as solid curves in each panel are the theoretically expected spectra for TM helices aligned perpendicular and parallel to the bilayer. The nearly identical spectra for GL_{12} and KL_{12} (Fig. 3A and B, respectively) show that the H-segments are predominately helical and aligned normal to the bilayer plane. The GL_{12} spectra adhere slightly less accurately to the ideal helix-normal spectra than the KL_{12} peptides, possibly indicating improved helicity due to the lysine residues. The spectra for GL_6 and KL_6 , on the other hand, are not readily interpretable; OCD spectra such as these have not been previously reported as far as we can establish and their interpretation is not obvious. The solution spectra (Fig. S2) of the $n = 6$ and 8 peptides indicate β -sheets for GL_n and random coil for KL_n , suggesting that the conformations may be more extended than in oriented bilayers. It may be significant that the OCD spectra have minima at about the same wavelength as for the solution CD spectra of GL_6 and GL_8 . One possibility is that the $n = 6$ and $n = 8$ peptides have a distorted, partially helical conformation. Although the OCD spectra provide no obvious information about the orientation of the peptides relative to the bilayer normal, one can say with confidence that the conformations and/or orientations of the peptides change as the lengths increase from $n = 6$ to $n = 12$ leucines.

We characterized the changes in the OCD spectra empirically by plotting the positions of wavelength minima as a function of n (Fig. 2B). The curves are remarkably similar to the probability curves of

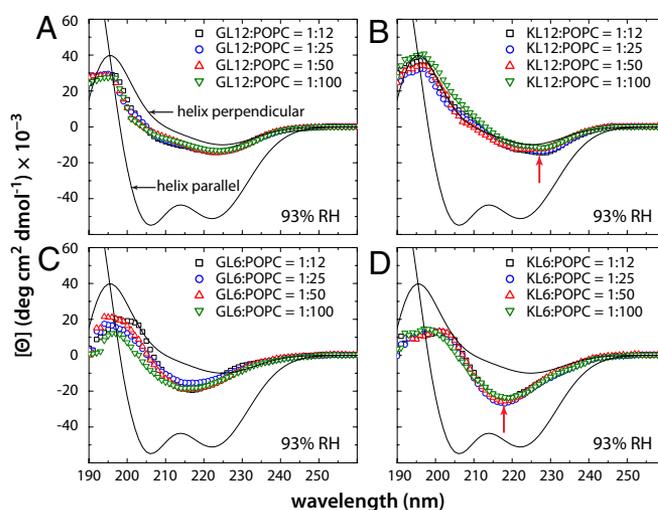


Fig. 3. Oriented circular dichroism (OCD) spectra of synthetic XXXX- L_n -XXXX peptides in oriented multilamellar POPC bilayer arrays for $X = G$ and $X = K$. Oriented multilamellar bilayer arrays were formed on 2.5-cm-diameter quartz plates, as described in detail elsewhere (9, 10, 27). Solid curves (labeled in A) represent theoretical spectra for transmembrane helices oriented normal and parallel to the membrane plane, as indicated. Red arrows in B and D indicate the positions of the minima in OCD curves. (A) Spectra for GL_{12} at various peptide/lipid mole ratios. The spectra indicate that the peptides are predominately α -helical and normal to the bilayer plane. (B) Spectra for KL_{12} at various peptide/lipid mole ratios. As for GL_{12} , the spectra are consistent with TM α -helices. (C and D) Spectra for GL_6 and KL_6 , respectively. GL_{12} : $X = G, n = 12$; GL_6 : $X = G, n = 6$; KL_{12} : $X = K, n = 12$; KL_6 : $X = K, n = 6$.

Fig. 2A in that they reveal a sharp break between $n = 8$ and $n = 10$. We cannot interpret this change structurally, but it is clear that for $n \geq 10$, the synthetic H-segments form TM α -helices, as predicted by the translocon experiments. The broad agreement between the synthetic-helix and translocon data are consistent with a common underlying physical principle.

Molecular Dynamics Simulations of Transmembrane Poly-leucine H-Segments. What is the conformation of the synthetic peptides in lipid bilayers, and how does the lipid bilayer adapt to them? To examine these questions, we carried out all-atom molecular dynamics simulations of the GL_n and KL_n peptides in POPC lipid bilayers for $n = 6, 8, 10, 12,$ and 20 (Figs. S4 and S5). The transbilayer phosphate-to-phosphate thickness of ER lipid membranes (12) is $\approx 38 \text{ \AA}$, similar to the thicknesses of DOPC and POPC at moderate hydrations (13). The number of atoms, the simulation box size ($a \times b \times c$) at equilibrium, the simulation time t_e (ns) required to reach equilibrium, and the lateral area per lipid (A_{lipid}) for each system are reported in Table 1. If it were technically feasible to run all-atom simulations for milliseconds or seconds, one might observe transitions between membrane-parallel interfacial helices and transmembrane helices. However, all-atom simulations on millisecond or longer time scales are not yet feasible. Nonetheless, our simulations provide critical insights into the behavior of bilayers in the vicinity of our H-segment transmembrane helices.

Our simulations (Table 1) indicate that A_{lipid} is proportional to the helical length: GL_{6-12} and KL_{6-12} all have much lower values than GL_{20} and KL_{20} . In the GL_n systems, A_{lipid} is directly proportional to the number of leucines, with a slope of 0.4 \AA^2 per residue, an intercept of 46.5 \AA^2 ($R = 0.996$). In the case of KL_n systems, on the other hand, it seems that $KL_8, KL_{10},$ and KL_{12} have reached a plateau of 51.4 \AA^2 . KL_6 does not respect this trend, but it also failed to remain helical. Moreover, A_{lipid} is 1.4 \AA^2 lower in the KL_{20} than in the GL_{20} system, which is likely explained by the fact that the KL_{20} system is 2 \AA thicker than the GL_{20} system

Table 1. Summary of the molecular dynamics simulations for various properties of GLn and KLn simulations in POPC bilayers

Peptide	n_{atoms}	$a \times b \times c, \text{\AA}^3$	t_{er}, ns	$A_{\text{lipid}}, \text{\AA}^2$	$n_{w10\text{\AA}}$	n_{bbN}	n_{bbO}	n_{scN}	Overlap	n_{HB}
GL6	62,913	$82.4 \times 83.8 \times 84.6$	27	49.3 ± 0.2	13	4.2	6.7	—	1.4	9.5
GL8	62,951	$84.0 \times 83.4 \times 83.4$	25	50.1 ± 0.2	8	4.5	7.9	—	2.0	10.4
GL10	62,989	$83.4 \times 85.9 \times 82.7$	23	50.6 ± 0.2	3	5.5	8.5	—	3.3	10.7
GL12	63,027	$86.6 \times 83.1 \times 81.5$	17	51.4 ± 0.2	3	7.3	10.9	—	3.8	14.4
GL20	63,179	$88.4 \times 87.5 \times 77.0$	17	55.2 ± 0.4	0	10.2	13.8	—	7.6	16.3
KL6	63,009	$83.9 \times 83.2 \times 83.9$	18	49.9 ± 0.2	6	4.1	7.9	11.9	0.9	23.0
KL8	63,047	$86.1 \times 83.8 \times 81.3$	7	51.5 ± 0.2	13	3.9	7.1	12.0	1.6	21.3
KL10	63,085	$86.7 \times 82.9 \times 81.5$	23	51.4 ± 0.2	11	4.9	10.3	11.8	2.3	24.8
KL12	63,123	$84.0 \times 85.5 \times 81.8$	11	51.3 ± 0.2	3	6.0	8.6	11.7	3.4	22.8
KL20	63,275	$84.8 \times 89.0 \times 79.0$	24	53.8 ± 0.3	1	11.1	14.5	12.7	6.6	31.7

The parameter n_{atoms} is the total number of atoms; a , b and c , the system size along the x -, y - and z -axes, respectively, averaged over the 10-nsec equilibrated run (the z -axis is the transmembrane axis); t_{er} , the time to reach equilibrium, i.e. the time required before the system reached a stable area ($a \times b$) for 10 consecutive ns; A_{lipid} , the mean area per lipid averaged over the 10-ns equilibrated run; $A_{\text{lipid}} = (a \times b)/(n_{\text{lipid}}/2)$, where $n_{\text{lipid}} = 280$. Various properties of Gn and Kn systems related to hydrogen-bonding are given by following tabulated values: $n_{w10\text{\AA}}$, the average number of water molecules within 10 Å of the center of the bilayer; n_{bbN} , n_{bbO} , and n_{scN} , are the total number of hydrogen-bonds involving backbone nitrogens as donors, backbone oxygens as acceptors, and side chain nitrogens as donors, respectively; overlap, the number of hydrogen-bonding interactions that are counted twice, because n_{bbN} and n_{bbO} sometimes include the same combination of donor and acceptor; n_{HB} , the total number of hydrogen-bonds, given by $n_{\text{bbN}} + n_{\text{bbO}} + n_{\text{scN}} - \text{overlap}$; —, no side-chain nitrogens.

as measured by the parameter c (Table 1). Overall, our results indicate that the systems thicken as the length of the inserted helices decrease. This observation seems contrary to any intuitive expectation that a bilayer (and consequently system thickness) should shrink with decreases in helix length to ensure hydrophobic matching between the two. There are 2 reasons for thickening. First, water plays an important stabilizing role. Second, the lipid bilayer adopts a distinct shape around each type of peptide. The roles of water and bilayer shape changes become apparent by comparing the GL6, KL6, GL12, and KL12 systems (Fig. 4).

The importance of water is revealed by snapshots of the 4 systems at the end of the 10-ns production run (Fig. 4 *Left*). Water (orange) penetrates deeply into the bilayer for GL6 and even through the bilayer for KL6, because of hydration of the peptide backbone. Water penetration is much less for the longer GL12 and KL12 peptides. This water penetration, shown quantitatively in Fig. 4 *Right*, results from the formation of cavities around each peptide due to the bilayer collapsing around the peptides (see below) through hydrophobic mismatch. Lysine groups are much longer and more polar than glycine, and these properties must necessarily have an important effect on a peptide's ability to stabilize itself through hydrogen-bonding interactions, as suggested by Fig. 4 and quantitated by the number of hydrogen bonds (n_{HB}) in Table 1. There are nearly twice as many hydrogen-bonding interactions for the KLn peptides, which means that the lysine residues add additional stability that backbone interactions alone cannot provide in the case of the GLn peptides. These lysine interactions are a complex form of snorkeling (14, 15) between lysine sidechains and lipid phosphate groups. A complete analysis of the hydrogen bonding of the peptides is included in Fig. S7.

The shape of the bilayer surrounding the GLn and KLn peptides reveals the adaptation of bilayer and peptide-to-hydrophobic mismatch (Fig. 5). The mismatch is made possible by changes in the lipid order parameters in the peptide neighborhood (Fig. S6). The shape of the bilayer around each peptide was analyzed by using a model based on the Landau-de Gennes theory, in which the free energy of lipid-protein interactions is determined by minimizing the functional integral of the local free energy density (16) (see *SI Text*). The variation of membrane thickness (z) as a function of the radial distance (r) from the peptide obtained from the shape analysis (Fig. 5) provides the entrée for parameterizing and computing the free energy change ΔG associated with mismatch deformation relative to GL20 and KL20 (see *SI Text*). In simple terms, the energetic cost of deformation increases as the square of the difference of the hydrophobic mismatch between the peptide and the unperturbed lipid bilayer. Qualitatively, Fig. 5 shows that

the GL12 and GL6 mismatch is higher than the mismatch between KL12 and KL6, meaning that it is energetically more costly to deform the bilayer around the GLn peptides than that around the KLn peptides. The computed relative free energies of deformation for GLn and KLn are compared in Fig. 2C. For values of n between 20 and 12, the mismatch energetic cost is about the same for the 2 peptide families. But for $n < 12$, the free energy cost for GLn is persistently greater than that for KLn. This result implies that it should be energetically easier for the translocon to insert short KLn peptides than to insert GLn peptides, which is in fact the case (Fig. 2A).

Discussion

We have used a consistent set of model TM segments (H-segments) for molecular biological, biophysical, and computational studies to gain insights into the mechanism of insertion of short poly-leucine segments across the ER membrane by the Sec61 translocon. We used a microsomal in vitro expression system to examine systematically the Sec61 insertion efficiency of short poly-leucine segments (L_n , $n = 4 \dots 12$) flanked at either end by 4-residue sequences of the form XXPX- L_n -XPXX with X = G, N, D, or K (Fig. 1). With the exception of X = K, insertion efficiency rises steeply between $n = 8$ to 12. For X = K, the range of steep rise shifts to shorter lengths ($n = 8$ to 10).

To explore the question of whether or not simple physical principles can explain the in vitro Sec61 results, we synthesized members of the GLn and KLn families of sequences, and examined by OCD their incorporation into POPC- and POPC:POPG-oriented multibilayers at various protein:lipid ratios and hydrations. The results were largely independent of lipid and hydration (Fig. S3). A comparison of the OCD data for GL6 and GL12 with KL6 and KL12 was revealing (Figs. 3 and 4). For $n \geq 10$, peptides incorporated into oriented lipid bilayers had OCD spectra unambiguously characteristic of transmembrane helices. This result shows that insertion of TM helices as short as 10 leucines by the Sec61 translocon are consistent with a physical partitioning of TM segments between translocon and lipid bilayer (5, 7).

OCD spectra for $n \leq 8$ were ambiguous with respect to orientation and secondary structure, but solution CD spectra suggested that the peptides are likely rich in β or extended conformations. A plot of the OCD wavelength minima against n (Fig. 4) revealed a sharp break between $n = 8$ and $n = 10$, mirroring the length-dependent Sec61 probability of insertion (Fig. 2). The solution spectra (Fig. S2) and the MD simulations (see below) raise the possibility that the shorter sequences cross the membrane with imperfect α -helical or extended conformations. But whatever their

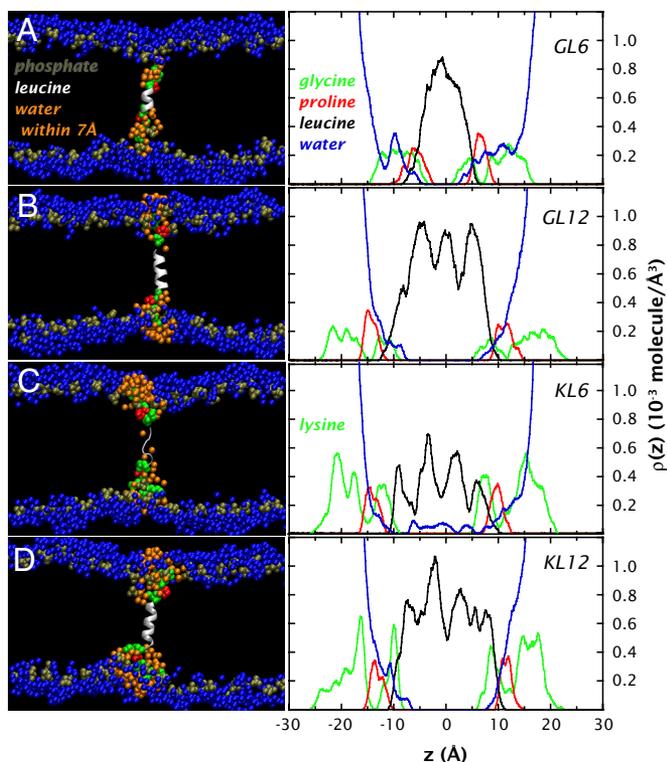


Fig. 4. Summary of all-atom molecular dynamics simulations of GL6, GL12, GK6, and GK12 peptides in POPC bilayers. (Left) Cut-away snapshots of the peptides spanning the lipid bilayer. The representative snapshots were taken during the simulation production runs (i.e., during the stable period after the simulation cell dimensions had reached steady values). Each POPC bilayer consisted of 280 POPC molecules (140 in each monolayer) and 8,400 water molecules. At the beginning of each simulation, the peptides were positioned at the bilayer center of mass with the leucines in an α -helical conformation and the XXPX...XPXX flanks in an extended conformation. Color code: gold, lipid phosphates; dark blue, waters associated with phosphates; white, leucine residues; green, glycine residues (A and B) or lysine residues (C and D); red, proline residues; orange, waters within 7 Å of the peptides. (Right) The production-run time-averaged transbilayer distributions of the waters and peptide amino acids for each simulation. The color code is the same as in Left. The atom densities are for heavy atoms, i.e., H atoms are not included. For reference, the oxygen atom number-density of bulk water is 33×10^{-3} atoms per \AA^3 . (A) Simulation of GL6 in a POPC bilayer. (B) Simulation of GL12 in a POPC bilayer. (C) Simulation of KL6 in a POPC bilayer. (D) Simulation of KL12 in a POPC bilayer.

conformations and orientations may be, one can conclude that the shorter sequences interact differently with bilayers than do the longer ones.

Molecular dynamics simulations (Fig. 5) show that the α -helical conformation is stable for all of the peptides examined, except for the KL6 peptide, which partially unfolds. The unfolding is stabilized in the simulation by waters entering the bilayer along the peptide backbone. We do not know, however, whether a similar phenomenon can occur in real membranes. The amount of hydrogen bonding with water and bilayer headgroups distinguishes the KL n from the GL n peptides (Table 1); the KL n peptides have about twice as many stabilizing hydrogen bonds due to the strong interactions of the positively charged lysine sidechains. This hydrogen bonding apparently allows the bilayer to adapt more readily to the KL n peptides, i.e., without so much hydrophobic mismatch, than to the GL n peptides. The resulting energetic cost of exposing the peptide bonds to the bilayer hydrocarbon interior (17) is apparently compensated for by improved interactions of the lysine sidechains with lipid phosphates and water (18) for KL6. Although one might expect a similar effect for KL8, none was observed, presumably

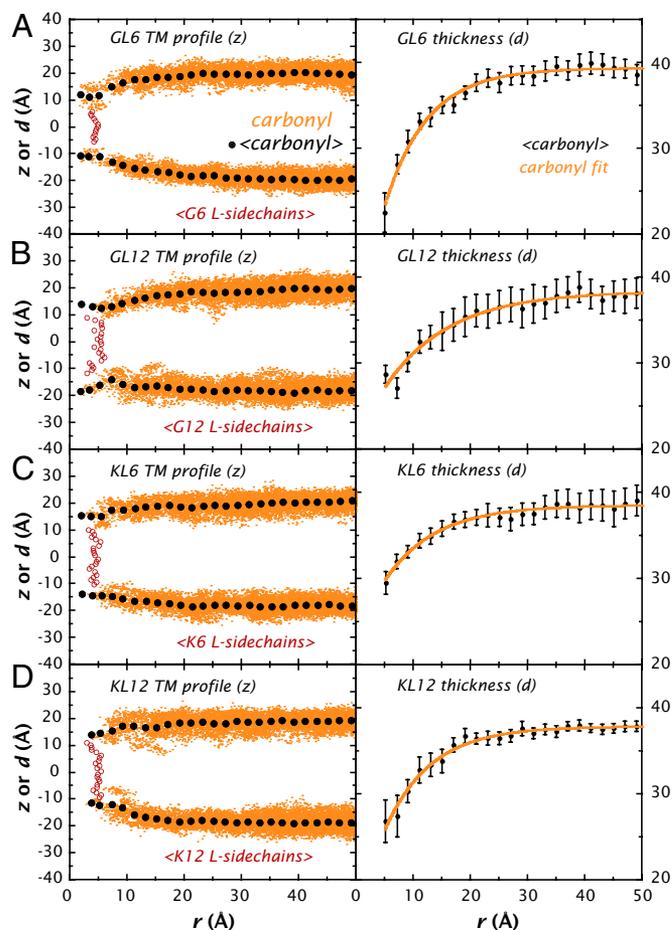


Fig. 5. Radial time-averaged profiles of the bilayers in the simulations. The average thickness $z(r)$ of the bilayer is shown as a function of the radial distance r from the average peptide axis. (Left) The time-averaged positions of the lipid carbonyl and peptide leucine groups. We ignored data for $r < 5$ Å because the average position of leucine sidechains is ≈ 5 Å. Taking periodic boundary conditions into consideration, we selected carbonyl groups within 50 Å of the center-of-mass of the helix, represented here by the orange dots. We then calculated the average position of 2-Å-thick rings of carbonyl groups centered on the helix's center of mass, indicated here by the black dots. Finally, we calculated the average position of leucine sidechains every 1 Å along the z -axis, represented here by the red circles. (Right) The computed average carbonyl-to-carbonyl separations across the bilayer at 2-Å intervals away from the helices (black points with error bars indicating the standard deviations of the means). The solid orange line represents the best-fit of the data to Eq. S1. A–D summarize the results of the analyses for GL6, GL12, KL6, and KL12 simulations, respectively.

because the cost of deforming the bilayer was smaller than the cost of exposing the peptide bonds to the bilayer interior.

An analysis of the shape of the bilayer in the vicinity of the peptides revealed consistent hydrophobic mismatch that, as expected, increased in proportion to the difference between the length of the leucine helices and the unperturbed bilayer thickness (Fig. 5). The shape analysis allowed the positive free energy cost of the hydrophobic mismatch for each peptide to be computed relative to the GL20 and KL20 peptides, which had essentially no mismatch (Fig. 2C). The energetic cost arises from the energy required to distort the bilayer to minimize the exposure of the nonpolar leucines to the aqueous phase. The most significant observation may be that the energetic costs for the GL n and KL n peptides rise roughly in parallel as n decreases to 10 or 12; but for $n < 10$, the energetic cost of mismatch becomes significantly smaller for the KL n peptides. This smaller energetic cost is most easily explained by the increased hydrogen bonding of the KL n peptides, which reduces, in effect, the amount of mismatch.

