Effects of flanking loops on membrane insertion of transmembrane helices: a role for peptide conformational equilibrium

Jian Gao and Jianhan Chen

Published Version Information

Citation: Gao, J., & Chen, J. (2013). Effects of flanking loops on membrane insertion of transmembrane helices: A role for peptide conformational equilibrium. The Journal of Physical Chemistry B, 117(28), 8330-8339.

Digital Object Identifier (DOI): doi:10.1021/jp402356c

Publisher's Link: http://pubs.acs.org/doi/full/10.1021/jp402356c

Effects of Flanking Loops on Membrane Insertion of Transmembrane Helices: A Role for Peptide Conformational Equilibrium

Jian Gao and Jianhan Chen*

Department of Biochemistry and Molecular Biophysics

Kansas State University, Manhattan, Kansas 66506

Runheads: Left: Gao and Chen
Right: Flanking loops on TM helix insertion

Submitted to *Journal of Physical Chemistry B* as an *Article*

Revised Version

*Corresponding Author: Phone: (785) 532-2518; Fax: (785) 532-7278; Email: jianhanc@ksu.edu

Abstract

The ability of a transmembrane helix (TMH) to insert into a lipid bilayer has been mainly understood based on the total hydrophobicity of the peptide sequence. Recently, Hedin et al. investigated the influence of flanking loops on membrane insertion of a set of marginally hydrophobic TMHs using translocon-based membrane integration assays. While the flanking loops were found to facilitate the insertion in most cases, counter examples also emerged where the flanking loops hinder membrane insertion and contradict the hydrophobicity and charge distribution analyses. Here, coarse-grained free energy calculations and atomistic simulations were performed to investigate the energetics and conformational details of the membrane insertion of two representative marginally hydrophobic TMHs with (NhaL and EmrL) and without (NhaA and EmrD) the flanking loops. The simulations fail to directly recapitulate the contrasting effects of the flanking loops for these two TMHs, due to systematic over-prediction of the stabilities of the transmembrane states that has also been consistently observed in previous studies. Nonetheless, detailed force decomposition and peptide conformation analyses suggest a novel mechanism on how the peptide conformational equilibrium in the aqueous phase may modulate the effects of flanking loops on membrane insertion. Specifically, the flanking loops in peptide EmrL interact strongly with the TMH segment and form stable compact conformations in the aqueous phase, which can hinder membrane absorption and insertion as these processes require extended conformations with minimal interactions between the flanking loops and TMH segment. This work also emphasizes the general importance of considering the peptide conformational equilibrium for understanding the mechanism and energetics of membrane insertion, an aspect that has not yet been sufficiently addressed in the literature.

Keywords: coarse-grained; molecular dynamics; PMF; transmembrane helices; hydrophobic scale

Introduction

Membrane proteins play key roles in biological processes ranging from signal transduction, molecular transport, drug binding, to cell communication and others[1]. They frequently contain one or more transmembrane helices (TMHs) with high fractions of hydrophobic residues[2]. Understanding the energetics and mechanism of membrane insertion of TMHs has remained an important topic in the study of membrane protein folding[3-5]. While some small single-spanning membrane proteins can insert spontaneously into membranes[6, 7], multispanning membrane proteins often contain TMHs with low hydrophobicity and need to be inserted with the aid of translocons[8-10]. It has been observed that a threshold of hydrophobicity appears to exist for TMHs in single-helix membrane proteins[11]. That is, the efficiency of membrane insertion is largely determined by the total hydrophobicity of all TMH residues, even though there exists a dependence on specific positioning of certain residues such as Trp and Tyr [8, 12].

It has also been recognized that the context of TMH can affect membrane insertion. For example, positive charges on the cytosolic side of a TMH can aid in its membrane insertion (i.e., the "positive-inside" rule)[13], even though the underlying molecular mechanism is unclear. Hedin et al. recently tested 16 "marginally hydrophobic" TMHs (mTMHs)[14] that were identified using a 'biological' hydrophobicity scale for translocon-mediated membrane insertion[11]. The apparent free energy (ΔG_{app}) of insertion into the endoplasmic reticulum (ER) membrane was measured using a Sec61 translocon-based membrane integration assay[14]. It was found that all these mTMHs inserted into the ER membrane with low efficiency and most had measured ΔG_{app} greater than 1.0 kcal/mol. Interestingly, including the adjacent N- and C-terminal loops favors the membrane insertion for all except two mTMHs investigated. For most peptides, the flanking loops increase the number of positively charged residues on the cytosolic side. Increased insertion efficiency could thus be at least partially explained by the positive-inside rule. However, such sequence-based hydrophobicity and charge characteristics analysis alone does not appear to be sufficient for explaining the effects of all flanking loops. For example, in the case where the flanking loops the most severely hinder translocon-mediated ER membrane insertion (EmrD TMH2 in Table 1), the insertion efficiency was reduced from 9% to 1%, even though more positive charges were introduced to the cytosolic side with the flanking loops and ΔG_{app} was predicted to remain similar using the translocon-based hydrophobicity scale [14].

Molecular simulations with different levels of detail can be explored to investigate the mechanism of how the flanking loops may affect membrane insertion of TMHs. Classical all-atom molecular mechanics[15-20], while providing atomistic details, are computationally demanding for the system size and time scale required for modeling membrane insertion processes. So-called coarse-gained (CG) models can dramatically extend to longer time and larger length scales by reducing the spatial resolution and/or adopting simplified pair-wise interaction schemes[21-23]. A particularly attractive CG approach is to map small groups of atoms to single pseudo-atoms (i.e., CG particles) and at the same time employs simplified effective interaction potentials[24]. Such CG models can preserve the microscopic features of both solute (protein) and solvent (water and membrane), which is essential to model nontrivial mutual responses of the protein and bilayer during membrane insertion. Importantly, with a modest level of coarse-graining (e.g., mapping 3-4 heavy atoms to one CG particle), quantitative and semi-quantitative agreement with atomistic simulations and experiments could be achieved. For example, the latest extension of the MARTINI lipid-water force field[25, 26] to proteins was carefully parameterized based on solvation and partition free energies to provide a good balance between peptide-solvent (including water and lipid) and solvent-solvent interactions[27, 28]. As such, the MARTINI force field and its various extensions to proteins have been remarkably effective in simulation of lipid self-assemblies as well as tmembrane insertion and interaction of various proteins[25-31].

In this work, we focus on two representative sets of mTMHs that were investigated by Hedin et al.[14] (see Table 1). In the case of Nha TMH4, the flanking loops significantly increase the insertion efficiency, whereas for EmrD TMH2 the flanking loops further reduce the insertion efficiency. CG simulations were first performed using the MARTINI force field[25, 26] to calculate the potentials of mean force (PMFs) of membrane insertion and to analyze the contributions of various inter and intra-molecular interactions to the overall free energy profiles. Atomistic simulations were then carried out to validate one of the key observations from the CG free energy simulations, that conformational equilibria of these peptides differ significantly in the aqueous phase and may play an important role in determining how the flanking loops may affect membrane insertion kinetics and thermodynamics. The current study also provides an opportunity to examine the efficacy of the MARTINI force field in mechanistic study of the interactions of rigid and flexible peptides with membrane bilayers.

Methods

Two mTMHs with and without flanking loops

We focused on two pairs of mTMHs in the current work. Table 1 lists the names, sequences, and measured translocon-mediated ER membrane insertion efficiencies for these peptides. The NhaA and NhaL pair represents a case where the flanking loops greatly enhance translocon-mediated ER membrane insertion efficiency (from 16% to 68%), and the EmrD and EmrL pair represents the opposite extreme of substantially depressed insertion efficiency with flanking loops (from 9% to 1%).

CG modeling: free energy simulation and analysis

CG molecular dynamics simulations were performed using a version of the MARTINI force field[25-27] that was recently implemented in CHARMM[32, 33]. An extensive set of liquid properties, and solvation free energies, and lipid/water partitioning free energies has been calculated to examine and establish the equivalency with the original GROMACS-based version (unpublished data). Per MARTINI convention, the backbone particle type was assigned based on the local secondary structure. The helical conformation of TMH was maintained by imposing harmonic restraints on the pseudo dihedral angles along the backbone particles with a force constant of 100 kcal/mol/radian^2. The flanking loops were presumed to lack stable secondary structures. The original MARTINI force field recommends assigning all backbone particles in the unstructured segments to type P5 (the most polar MARTINI particle type) except for alanines (P4) and prolines (Na). However, analysis of the PMFs of interactions between various particle types (see Supporting Materials Fig. S1) shows that an assignment of P5 overestimates the hydrophilicity of peptide backbone in the coil state. In particular, the P5-P5 interaction is weaker than either P5-P4 (backbone-water) or P4-P4 (water-water) interaction (Fig. S1). This is not consistent with previous calculations using atomistic force fields[34, 35]. As such, the backbone particle type for non-proline amino acids in the flanking loops was assigned P2, which represents a weaker polar particle type and provides more favorable peptide-peptide interactions (Fig. S1). Necessary counter ions for neutralizing the systems were modeled by particle type “Qa” (anions) and “Qd” (cations).

A pre-equilibrated bilayer with 256 CG dipalmitoylphosphatidylcholine (DPPC) molecules was in all free energy calculations. The final simulation box was ~8.8 × 8.8 × 14.0 nm^3 for NhaA and EmrD, and ~8.8 × 8.8 × 18.5 nm^3 for NhaL and EmrL. Umbrella sampling was used to calculate the PMF of membrane insertion. Harmonic restraint potentials were placed on the center of mass of the peptide, centered at every 1.0 Å along the membrane normal (z axis) with a force constant 2.39 kcal/mol/$Å^2$. The total number of umbrella sampling windows was 171 (-85 to 85 Å) and 121 (-60 to 60 Å) for peptides with and without loops, respectively. For preparation of the initial conformation for each window, we first relaxed the peptide in water for 100 ns. The final peptide structure was then inserted into the pre-equilibrated and solvated DPPC model bilayer at the targeted positions. The TMH segment was aligned with the membrane normal with the N-terminus pointing up in all initial configurations. The numbers of lipid molecules in the top and bottom leaflets were checked, which did not show large differences. The initial configurations were equilibrated for 2 ns before 100 ns production simulations were performed under atmospheric pressure and at 323 K. The integration time step was 20 fs and the conformations were saved every 2 ps. The dielectric constant was set at 15. Long-range electrostatic and van der Waals interactions were shifted and then gradually switch off from 9.0 to 14.0 Å.

Unbiased PMFs were calculated using the weighted histogram analysis method (WHAM)[36]. The final PMFs appear to be well converged (Fig. S2) and show even baselines above and below the membrane (Fig. 1). Thus, no symmetrization operation was applied[37]. Force decomposition[38] was performed to further analyze the contributions of interactions between various moieties to membrane insertion. For this, the peptides were divided into three segments including the TMH and two flanking loops; the solvents were grouped into water, lipid-head, and lipid-tail portions. Accordingly, a total number of nine decomposed PMFs were extracted for each peptide. In contrast to the total PMFs, many of the decomposed components display flat but uneven baselines, especially for NhaL and EmrL. This is most likely due to incomplete sampling of the peptide conformation with flexible loops, especially when in contact with the membrane. For this, simple linear correction was applied to remove the baseline offsets[39].

Atomistic explicti water and membrane simulations

The MARTINI force field was not parameterized with a focus on the ability to describe the peptide conformational equilibrium[27]. Atomistic simulations were thus performed in explicit

solvent to validate the differences in conformations of peptides with flanking loops observed in the CG simulations. For both NhaL and EmrL, we first generated initial conformations that contained ideal helices in the predicted TMH regions and fully extended loops. These initial conformations were relaxed using 10 ns simulations in the GBSW implicit solvent[35, 40], before solvated in TIP3P water boxes of appropriate sizes. The final solvated systems, shown in Fig. S3, contain 9229 TIP3P waters for EmrL and 7707 TIP3P waters for NhaL. The CHARMM22/CMAP force field was used[41-43]. Periodic boundary conditions were imposed, and the particle mesh Ewald method was used for long-range electrostatic interactions[44]. The van de Waals interactions were smoothly switched off from 12 to 13 Å. Lengths of all hydrogen-related bonds were kept constant using SHAKE[45], and the MD time step was 2 fs. After 100 ps of equilibration, 50 ns production simulations were carried out for both systems at 323 K and under atmospheric pressure. Significant conformational re-arrangements were observed during the initial stages of production simulations, and the final histograms were calculated using the last 30 ns trajectories.

Two additional atomistic simulations were performed for both NhaL and EmrL in the interfacial-like and TM states. For this, two more extended conformations sampled during the GBSW equilibration simulations (see above) were first selected and then placed in proper configurations with respect to a pre-equilibrated DPPC bilayer (see Fig. S3). The bilayer consists of 256 lipids and was generated using the CHARMM-GUI membrane builder[46]. In the interfacial configurations, the peptides were placed horizontally in orientations that allowed extensive direct contacts with the lipid head groups. For the TM configurations, the peptides were placed along the membrane normal and with the center-of-mass of the TMH segments at the membrane center. Proper counter ions were added to neutralize the total charge. The final explicit water/membrane systems contain 16486 and 19897 water molecules for NhaL in the interfacial and TM states, respectively; and 16451 and 18520 water molecules for EmrL in the interfacial and TM states, respectively. The dimensions of the simulation boxes were approximately 9 nm × 9 nm × 10 nm for the IF configurations and 9 nm × 9 nm × 11 nm for the TM configurations. A protocol similar to the one used our previous works[47, 48] was applied to equilibrate the system at 323 K using CHARMM[32, 33], during which various restraints were applied to selected heavy atoms and with gradually decreasing restraint strengths. The same non-bonded options were the same as those used in the aqueous simulations (see above). The equilibrated structures were then

used to initiate production simulations using NAMD[49]. Another 100 ps equilibration simulation with weak harmonic positional restraints on protein backbone atoms (k = 0.1 kcal/mol/Å^2) was used to suppress potential strains during subtle differences between CHARMM and NAMD. Nonbonded options equivalent or identical to those used in CHARMM equilibration simulations were used, and the length of the final NAMD production simulations was 30 ns for all four systems. Only conformations sampled during the last 20 ns were included in the final analysis. VMD[50] was used for preparation of all molecular images presented in this work.

Results and Discussion

Convergence of overall free energy profiles and comparison to translocon-based measurements

As shown in Fig. S2, the overall PMFs of membrane insertion appears to be well converged. The PMFs of both NhaL and EmrL insertion show minimal variance when calculated using different subsets of the umbrella sampling data. For example, the maximal deviation between PMFs calculated using the last 20 ns and full 100 ns is less than 2 kcal/mol, and much smaller for PMFs of NhaA and EmrD without long flexible loops. As such, all subsequent analysis was preformed by including the last 40 ns of the umbrella sampling data.

The final PMFs of membrane insertion, shown in Fig. 1, predict both mTMHs to be very stable in the POPC membrane. The stability of the TM state (taken at z = 0) is about -16.0 kcal/mol for NhaA and -8.9 kcal/mol for EmrD with respect to the aqueous state (taken at the largest distances from the bilayers). While the prediction that EmrD is less stable than NhaA in the TM state is correct, the actual stability values for both peptides are much larger than the translocon results [14] (+0.97 kcal/mol for NhaA and +1.3 kcal/mol for EmrD; see Table 1). We note that the free energy associated with TMH helix folding is not included in the PMFs calculated, as the peptides were restrained to maintain helical conformation in the TM regions throughout the insertion process. However, such contribution to the overall PMF should be no greater than a few kcal/mol. For example, the helical restraints should increase the relative stability of the (helical) TM state by ~1.5 kcal/mol, if the TMH segment only retains ~10% residual helicity in the aqueous state. Importantly, similar large apparent differences in the magnitudes of energetics of membrane insertion between translocon-based measurements and molecular simulations

(atomistic or coarse-grained) have been consistently observed in previous studies[51-54]. While inherent imperfections in the force fields (and the free energy calculation protocols) are clearly contributing factors, the order of magnitude discrepancy has been mainly attributed to differences in lipid compositions and protein concentrations of ER and model membranes as well as interactions with translocon machinery. It has also been proposed that the interfacial state might be used instead of the aqueous state as the reference for comparison with the translocon data[5]. Indeed, the stabilities of the TM states would be revised to -1.5 kcal/mol for NhaA and +5.9 kcal/mol for EmrD in reference to the local minima near z = -1.6 nm. The revised values are closer to the experimental results, especially if including the estimated 1.5 kcal/mol stability over estimation introduced by helical restraints (see above). Nonetheless, it has been shown that, despite the order of magnitude differences, there exists a strong correlation between translocon-based hydrophobicity scale and partition free energies derived from simulations or other experimental approaches[5, 55]. Therefore, the free energy simulations should provide useful insights for understanding the observed effects of flanking loops on membrane insertion.

Effects of the flanking loops on the overall free energy profiles

As shown in Fig. 1, the inclusion of flanking loops leads to significant changes in the overall free energy profiles for both peptides. The stabilities of the TM states are significantly increased, to ~ -31.5 kcal/mol (from -16.0 kcal/mol) for NhaL and -28.5 kcal/mol (from -8.9 kcal/mol) for EmrL. Detailed molecular analysis suggests that in the TM states the helical segments are similar with and without the flanking loops (e.g., see Figs. 2 and 3), as expected. In particular, the tilt angles of the TM helices remain similar (Fig. 4), and the flanking loops remain at the membrane/water interface (Fig. 3). While the stabilizing effects of the flanking loops of the NhaL peptide appear consistent with the experimental result[14] (Table 1), the predictions that the flanking loops are stabilizing for both NhaL and EmrL is clearly incorrect. However, it is unlikely that the apparent over-predication of the stabilizing effects of flanking loops is mainly due to systematic artifacts in the MARTINI force field[26, 27]. The reason is that the MARTINI force field has been specifically calibrated to provide reasonable balance between peptide-water and peptide-lipid interactions, based on the water/oil and water/bilayer partition free energies and PMF of membrane insertion of side chain analogs[27]. A recent study[56] further shows that the partitioning of a series of pentapeptides (Ac-WLXLL) at the cyclohexane/water and POPC/water

interface calculated using the MARTINI force field agrees well with the experimental Wimley-White hydrophobic scale[57] (except for charged residues). Therefore, over-prediction of the TM stabilities of NhaL and EmrL is most likely due to the same factors that underlie the order of magnitude discrepancy between calculated and translocon-derived stabilities of other TMH without flanking loops (see the previous section). Another significant effect of the flanking loops is that the interfacial minima (~ ±2 nm) are no longer present. The snapshots of membrane insertion (Figs. 2 and 3) illustrate that long flanking loops strongly perturb the interfacial bound states, not only disrupting the interactions of the peptides with the lipid molecules but also inducing substantial disorder in the local bilayer structure (e.g., Fig. 3B, at -2.8 nm).

A potential role of peptide conformational equilibrium in insertion efficiency

It is disappointing that the simulations have consistently over predicted the TM stabilities (in the current work and others) such that the overall PMF profiles fail to directly recapitulate the contrasting effects of flanking loops on the insertion efficiency of NhaL and EmrL. Nonetheless, a few detailed observations together suggest a potential role of peptide conformational equilibria in determining the effects of flanking loops in membrane insertion. The first indication is the observation that the flanking loops significantly elevate the free energy barriers for membrane insertion. As shown in Fig. 1, the barrier of insertion increases from about 5 kcal/mol to ~8 kcal/mol for NhaL (near z = 5 nm) and over 13 kcal/mol for EmrL (near z ~ -4 nm). Analysis of the peptide conformational equilibrium during insertion (e.g., see Figs. 2 and 3) suggest that the free energy barriers likely arise from the need for the peptides to adopt more extended conformations for absorption to the membrane interface and for inserting into the membrane. Indeed, as shown in Fig. 5, the average radii of gyration (R_g) significantly increase when the peptides transfer from the aqueous phase to the membrane bound states (green traces). Importantly, the magnitude of peptide size change upon membrane insertion appears to correlate well with the free energy barrier height. For example, the peptide size increase is larger with the flanking loops, and the largest size increase is required for the insertion of EmrL. Peptide conformational transition as the origin of the insertion free energy barriers is also supported by force decomposition analysis (see below). The free energy penalty associated with the peptide conformational changes likely also contribute to the lack of interfacial minima (see above).

While the insertion efficiency measured from the translocon-based essay should mainly capture the thermodynamic aspects[1], it is plausible that substantial barriers (on the order of ~10 kcal/mol) can significantly reduce the apparent efficiency of membrane integration. Comparing all peptides simulated, it is remarkable that EmrL has similar R_g in the aqueous phase compared to much shorter EmrD (Fig. 5B), while NhaL is substantially more extended compared to NhaA (Fig. 5A). The stronger ability of the flanking loops in EmrL to pack more tightly with the TMH segment should also have direct thermodynamic consequences: it will preferentially stabilize in the aqueous phase over the membrane bound states. The notion that peptide conformational equilibrium in the aqueous phase plays an important role in governing the membrane insertion efficiency is a key result of the current study. To the best of our knowledge, this has not been discussed previously.

PMF decompositions analysis

PMF decomposition analysis was performed to further dissect the contributions of interactions between various peptide segments and solvent components to the overall free energy profiles. The contributions from the interactions between the two flanking loops and water and lipid head/tail groups are summarized in Fig. 6, and all decomposed components are provided in the Supporting Materials Fig. S4. The analysis shows that the interactions between the flanking loops and lipid molecules are stabilizing for all four peptides, and desolvation is the main force that opposes membrane insertion (Fig. 6A and D; Fig. S4B and F). Nonetheless, the total contribution of intermolecular interactions between the peptides and the solvent (water + lipid molecules) strongly favors the membrane bound states, particularly the TM state (Fig. S4A). Interestingly, the total contributions from peptide-solvent interactions do not contribute to the free energy barriers of insertion present in the overall PMF profiles shown in Fig. 1. This further supports the proposed roles of peptide conformational equilibria in modulating the membrane integration efficiency. The decomposition analysis also reveals that the TMH segment of EmrD and EmrL is intrinsically unstable in the TM state, with the net contribution from TMH-solvent interactions to be destabilizing (Fig. S4E). As such, the short loops (GGPG and GPGG) actually play key stabilizing roles in membrane insertion of NhaA and EmrD. Intermolecular interactions of the long flanking loops in NhaL and EmrL with the solvent molecules provide similarly strong

stabilizing effects on the TM states (Fig. S4I and M), which appears consistent with the hydrophobicity scale and charge distribution analysis[14].

Conformational properties of NhaL and EmrL in all-atom explicit water and bilayer

The predicted compact state of EmrL in the aqueous phase provides a key clue for rationalizing the destabilizing effects of the flanking loop. As the current MARTINI force field has not been specifically parameterized to properly describe flexible proteins[27, 58], atomistic simulations were performed to further validate the conformational properties of NhaL and EmrL in the aqueous, interfacial-like and TM states. The final snapshots of these simulations are shown in Fig. 7. Interestingly, the interfacial-like states were unstable for both NhaL and EmrL (even though care was taken to place the peptides in configurations that allowed substantial peptide-lipid contacts; see Fig. S3). For NhaL, the peptide underwent partial dissociation from the membrane rapidly during the simulation and only remained in contact with the bilayer through the N-terminus (see Fig. S5). For EmrL, the peptide was fully dissociated from the bilayer around the 10^{th} ns, adopted compact conformations while diffusing freely in water, and the formed partial contacts with the bottom leaflet of the bilayer (due to periodic boundary conditions) at the end of the simulation (see Fig. S6). The observed instability of interfacial-like configurations appears to be highly consistent with the presence of strong membrane absorption barriers predicted from the MARTINI simulations (see Fig. 1).

Fig. 7 compares the size distributions of NhaL and Emrl sampled during these atomistic simulations. These R_g distributions are in quantitative agreement with those derived from the MARTINI simulations (Fig. 5). In particular, EmrL adopts highly compact conformations in water and leads to a rather narrow R_g distribution that peaks at ~ 1.1 nm; whereas NhaL is significantly more extended and more flexible in water with an average R_g ~ 1.6 nm (solid traces in Fig. 8). In contrast, both the interfacial-like and TM states require significant expansion of the peptides, to R_g > 1.8 nm for both NhaL and EmrL. Note that EmrL did not sample a true interfacial state during the 30-ns simulation (see Fig. S6) and the calculated R_g distribution appears to largely reflect the aqueous conformations. Together, the atomistic simulations further support the notion that the self-interactions between the flanking loops and TMH segment in EmrL could impede the membrane association and insertion, as the later would require the peptide to adopt much more extended conformations.

Conclusions

Free energy simulations were performed using the MARTINI force field to understand the contrasting effects of flanking loops in the membrane insertion efficiencies of two marginally stable transmembrane helices previously investigated using the translocon-based assay[14]. The hydrophobicity scale predicts minimal effect of the flanking loops on the insertion efficiencies of both NhaA and EmrD. Intriguingly, the translocon-based membrane integration assay showed that NhaL inserted about 5 times more efficiently compared to NHaA, while EmrL inserted about 9 times less efficiently with the flanking loop[14]. The calculated overall PMF profiles, despite very good apparent convergence, systematically over-predict the stabilities of the TM states compared to the translocon results, by about an order of magnitude. Similar over-predictions have also been consistently observed in previous studies, and are frequently attributed to key characteristics of translocon-based assays that are not fully represented in simulations[51-54].

Nonetheless, further analysis of the molecular detail and force decomposition suggests a novel mechanism on how the peptide conformational equilibrium in the aqueous phase may modulate the effects of flanking loops on membrane integration. Specifically, the force decomposition analysis shows that intermolecular interactions of the flanking loops with the solvent molecules provide similarly stabilizing effects for the TM states, which is apparently consistent with the hydrophobicity scale analysis[14]. Thus, intermolecular interactions or the hydrophobicity scale alone is not sufficient to explain the contrasting effects of flanking loops on insertion of NhaA and EmrD. Instead, whether the flanking loops may promote or inhibit membrane insertion also depends on how they interact with the TMH segment in the aqueous phase. The simulations reveal that the flanking loops in EmrL interact strongly with the TMH segment and lead to stable, compact conformations in the aqueous phase. This prediction has been confirmed by atomic simulations in explicit solvent. A stabilized compact aqueous state can impede membrane absorption and insertion, both kinetically and thermodynamically, as the later requires the peptide to adopt open conformations with minimal interactions between the TMH and flanking loops. This provides a molecular mechanism for explaining the contrasting effects of flanking loops on insertion of NhaA and EmrD. The current work also emphasizes the importance of considering the peptide conformational equilibrium in the aqueous phase in understanding the

membrane insertion mechanism and thermodynamics in general, an important aspect that has not attracted sufficient attentions.

Acknowledgements

This work was supported by the National Science Foundation (MCB 0952514) and the National Institutes of Health (GM074096). Part of the computing for this project was performed on the Beocat Research Cluster at Kansas State University**,** which is funded in part by NSF grants CNS-1006860, EPS-1006860, and EPS-0919443. This work is contribution number 13-287-J from the Kansas Agricultural Experiment Station.

Supporting Information Available

The PMFs of pair-wise interactions between various MARTININ polar particle types, total PMFs of membrane insertion calculated using different simulation lengths, force decomposition analysis, and various snapshots of the atomistic simulations of NhaL and EmrL in explicit water and membrane. This information is available free of charge via the Internet at http://pubs.acs.org/.

REFERENCES

1. White, S. H.; Wimley, W. C., Membrane Protein Folding and Stability: Physical Principles. *Annu. Rev. Biophys. Biomol. Struct.* **1999,** *28*, 319-365.

2. Bowie, J. U., Solving the Membrane Protein Folding Problem. *Nature* **2005,** *438*, 581-589.

3. Popot, J. L.; Engelman, D. M., Helical Membrane Protein Folding, Stability, and Evolution. *Annu. Rev. Biochem.* **2000,** *69*, 881-922.

4. Im, W.; Brooks, C. L., Interfacial Folding and Membrane Insertion of Designed Peptides Studied by Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 6771-6776.

5. Chetwynd, A.; Wee, C. L.; Hall, B. A.; Sansom, M. S., The Energetics of Transmembrane Helix Insertion into a Lipid Bilayer. *Biophys. J.* **2010,** *99*, 2534-40.

6. Dempsey, C. E., The Actions of Melittin on Membranes. *Biochim. Biophys. Acta* **1990,** *1031*, 143-161.

7. Herrera, A. I.; Al-Rawi, A.; Cook, G. A.; Gao, J.; Iwamoto, T.; Prakash, O.; Tomich, J. M.; Chen, J., Structural Characterization of Two Pore-Forming Peptides: Consequences of Introducing a C-Terminal Tryptophan. *Proteins* **2010,** *78*, 2238-50.

8. Hessa, T.; Kim, H.; Bihlmaier, K.; Lundin, C.; Boekel, J.; Andersson, H.; Nilsson, I.; White, S. H.; von Heijne, G., Recognition of Transmembrane Helices by the Endoplasmic Reticulum Translocon. *Nature* **2005,** *433*, 377-81.

9. White, S. H.; von Heijne, G., Transmembrane Helices before, During, and after Insertion. *Curr. Opin. Struct. Biol.* **2005,** *15*, 378-386.

10. Blobel, G.; Dobberstein, B., Transfer of Proteins across Membranes. I. Presence of Proteolytically Processed and Unprocessed Nascent Immunoglobulin Light Chains on Membrane-Bound Ribosomes of Murine Myeloma. *J. Cell Biol.* **1975,** *67*, 835-51.

11. Hessa, T.; Meindl-Beinker, N. M.; Bernsel, A.; Kim, H.; Sato, Y.; Lerch-Bader, M.; Nilsson, I.; White, S. H.; von Heijne, G., Molecular Code for Transmembrane-Helix Recognition by the Sec61 Translocon. *Nature* **2007,** *450*, 1026-30.

12. Ulmschneider, M. B.; Sansom, M. S.; Di Nola, A., Properties of Integral Membrane Protein Structures: Derivation of an Implicit Membrane Potential. *Proteins* **2005,** *59*, 252-65.

13. Lerch-Bader, M.; Lundin, C.; Kim, H.; Nilsson, I.; von Heijne, G., Contribution of Positively Charged Flanking Residues to the Insertion of Transmembrane Helices into the Endoplasmic Reticulum. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105*, 4127-32.

14. Hedin, L. E.; Ojemalm, K.; Bernsel, A.; Hennerdal, A.; Illergard, K.; Enquist, K.; Kauko, A.; Cristobal, S.; von Heijne, G.; Lerch-Bader, M.; Nilsson, I.; Elofsson, A., Membrane Insertion of Marginally Hydrophobic Transmembrane Helices Depends on Sequence Context. *J. Mol. Biol.* **2010,** *396*, 221-9.

15. Roux, B., Ion Conduction and Selectivity in K+ Channels. *Annu. Rev. Biophys. Biomol. Struct.* **2005,** *34*, 153-171.

16. Mackerell, A. D., Empirical Force Fields for Biological Macromolecules: Overview and Issues. *J. Comput. Chem.* **2004,** *25*, 1584-1604.

17. Feig, M., Implicit Membrane Models for Membrane Protein Simulation. *Methods Mol. Biol.* **2008,** *443*, 181-96.

18. Im, W.; Feig, M.; Brooks, C. L., 3rd, An Implicit Membrane Generalized Born Theory for the Study of Structure, Stability, and Interactions of Membrane Proteins. *Biophys. J.* **2003,** *85*, 2900-18.

19. Panahi, A.; Feig, M., Dynamic Heterogeneous Dielectric Generalized Born (Dhdgb): An Implicit Membrane Model with a Dynamically Varying Bilayer Thickness. *J. Chem. Theory Comput.* **2013,** *9*, 1709-1719.

20. Domene, C.; Bond, P. J.; Sansom, M. S., Membrane Protein Simulations: Ion Channels and Bacterial Outer Membrane Proteins. *Adv. Protein Chem.* **2003,** *66*, 159-93.

21. Ayton, G. S.; Noid, W. G.; Voth, G. A., Multiscale Modeling of Biomolecular Systems: In Serial and in Parallel. *Curr. Opin. Struct. Biol.* **2007,** *17*, 192-198.

22. Klein, M. L.; Shinoda, W., Large-Scale Molecular Dynamics Simulations of Self-Assembling Systems. *Science* **2008,** *321*, 798-800.

23. Sherwood, P.; Brooks, B. R.; Sansom, M. S. P., Multiscale Methods for Macromolecular Simulations. *Curr. Opin. Struct. Biol.* **2008,** *18*, 630-640.

24. Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L., A Coarse Grain Model for Phospholipid Simulations. *J. Phys. Chem. B* **2001,** *105*, 4464-4470.

25. Marrink, S. J.; de Vries, A. H.; Mark, A. E., Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **2004,** *108*, 750-760.

26. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H., The Martini Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007,** *111*, 7812-7824.

27. Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J., The Martini Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008,** *4*, 819-834.

28. Bond, P. J.; Sansom, M. S. P., Insertion and Assembly of Membrane Proteins Via Simulation. *J. Am. Chem. Soc.* **2006,** *128*, 2697-2704.

29. Shih, A. Y.; Arkhipov, A.; Freddolino, P. L.; Schulten, K., Coarse Grained Protein-Lipid Model with Application to Lipoprotein Particles. *J. Phys. Chem. B* **2006,** *110*, 3674-3684.

30. Chetwynd, A. P.; Scott, K. A.; Mokrab, Y.; Sansom, M. S. P., Cgdb: A Database of Membrane Protein/Lipid Interactions by Coarse-Grained Molecular Dynamics Simulations. *Mol. Membr. Biol.* **2008,** *25*, 662-669.

31. Marrink, S. J.; de Vries, A. H.; Tieleman, D. P., Lipids on the Move: Simulations of Membrane Pores, Domains, Stalks and Curves. *Biochimica Et Biophysica Acta-Biomembranes* **2009,** *1788*, 149-168.

32. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M., Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983,** *4*, 187-217.

33. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al., Charmm: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009,** *30*, 1545-1614.

34. Im, W.; Chen, J.; Brooks, C. L., 3rd, Peptide and Protein Folding and Conformational Equilibria: Theoretical Treatment of Electrostatics and Hydrogen Bonding with Implicit Solvent Models. *Adv. Protein Chem.* **2005,** *72*, 173-98.

35. Chen, J. H.; Im, W. P.; Brooks, C. L., Balancing Solvation and Intramolecular Interactions: Toward a Consistent Generalized Born Force Field. *J. Am. Chem. Soc.* **2006,** *128*, 3728-3736.

36. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M., The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules .1. The Method. *J. Comput. Chem.* **1992,** *13*, 1011-1021.

37. Roux, B.; Allen, T.; Berneche, S.; Im, W., Theoretical and Computational Models of Biological Ion Channels. *Q. Rev. Biophys.* **2004,** *37*, 15-103.

38. Roux, B.; Karplus, M., Ion Transport in a Model Gramicidin Channel: Structure and Thermodynamics. *Biophys. J.* **1991,** *59*, 961-981.

39. Allen, T. W.; Andersen, O. S.; Roux, B., Molecular Dynamics - Potential of Mean Force Calculations as a Tool for Understanding Ion Permeation and Selectivity in Narrow Channels. *Biophys. Chem.* **2006,** *124*, 251-267.

40. Im, W. P.; Lee, M. S.; Brooks, C. L., Generalized Born Model with a Simple Smoothing Function. *J. Comput. Chem.* **2003,** *24*, 1691-1702.

41. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al., All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998,** *102*, 3586-3616.

42. Mackerell, A. D.; Feig, M.; Brooks, C. L., Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004,** *25*, 1400-1415.

43. MacKerell, A. D.; Feig, M.; Brooks, C. L., Improved Treatment of the Protein Backbone in Empirical Force Fields. *J. Am. Chem. Soc.* **2004,** *126*, 698-699.

44. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993,** *98*, 10089-10092.

45. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* **1977,** *23*, 327-341.

46. Jo, S.; Kim, T.; Im, W., Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLoS One* **2007,** *2*, e880.

47. Herrera, H. I.; Al-Rawi, A.; Cook, G. A.; Gao, J.; Iwamoto, T.; Prakash, O.; Tomich, J. M.; Chen, J., Structural Characterization of Two Channel-Forming Peptides: Consequences of Introduction a C-Terminal Tryptophan. *Proteins: Struct., Funct., Genet.* **2010,** *78*, 2238-2250.

48. Bukovnik, U.; Gao, J.; Cook, G. A.; Shank, L. P.; Seabra, M. B.; Schultz, B. D.; Iwamoto, T.; Chen, J.; Tomich, J. M., Structural and Biophysical Properties of a Synthetic Channel-Forming Peptide: Designing a Clinically Relevant Anion Selective Pore. *Biochim. Biophys. Acta* **2012,** *1818*, 1039-48.

49. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kal, L.; Schulten, K., Scalable Molecular Dynamics with Namd. *J. Comput. Chem.* **2005,** *26*, 1781-1802.

50. Humphrey, W.; Dalke, A.; Schulten, K., Vmd: Visual Molecular Dynamics. *J. Mol. Graph.* **1996,** *14*, 33-&.

51. Rychkova, A.; Vicatos, S.; Warshel, A., On the Energetics of Translocon-Assisted Insertion of Charged Transmembrane Helices into Membranes. *Proc Natl Acad Sci U S A* **2010,** *107*, 17598-603.

52. Bond, P. J.; Wee, C. L.; Sansom, M. S., Coarse-Grained Molecular Dynamics Simulations of the Energetics of Helix Insertion into a Lipid Bilayer. *Biochemistry (Mosc.)* **2008,** *47*, 11321-31.

53. Gkeka, P.; Sarkisov, L., Interactions of Phospholipid Bilayers with Several Classes of Amphiphilic Alpha-Helical Peptides: Insights from Coarse-Grained Molecular Dynamics Simulations. *J. Phys. Chem. B* **2010,** *114*, 826-39.

54. Gumbart, J.; Chipot, C.; Schulten, K., Free-Energy Cost for Translocon-Assisted Insertion of Membrane Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108*, 3596-601.

55. Vostrikov, V. V.; Hall, B. A.; Greathouse, D. V.; Koeppe, R. E., 2nd; Sansom, M. S., Changes in Transmembrane Helix Alignment by Arginine Residues Revealed by Solid-State Nmr Experiments and Coarse-Grained Md Simulations. *J. Am. Chem. Soc.* **2010,** *132*, 5803-11.

56. Singh, G.; Tieleman, D. P., Using the Wimley–White Hydrophobicity Scale as a Direct Quantitative Test of Force Fields: The Martini Coarse-Grained Model. *J. Chem. Theory Comput.* **2011,** *7*, 2316-2324.

57. Wimley, W. C.; White, S. H., Experimentally Determined Hydrophobicity Scale for Proteins at Membrane Interfaces. *Nat. Struct. Biol.* **1996,** *3*, 842-848.

58. Seo, M.; Rauscher, S.; Pomes, R.; Tieleman, D. P., Improving Internal Peptide Dynamics in the Coarse-Grained Martini Model: Toward Large-Scale Simulations of Amyloid- and Elastin-Like Peptides. *J. Chem. Theory Comput.* **2012,** *8*, 1774-1785.

Tables

TABLE 1. Names, sequences and translocon-mediated ER Membrane insertion efficiencies of two mTMHs with and without flanking loops.

Short Name	Protein, TM helix (length) [a]	Sequence [b]	Insertion efficiency (%) [c]	ΔG_{app} (kcal/mol)
NhaA	NhaA, TMH4 (28)	*GGPG* WAIPA ATDIA FALGV LALLG *GPGG*	16	0.97
NhaL	NhaL, TMH4F (47)	*GGPG* ADPIT **R**EG WAIPA ATDIA FALGV LALLG S**R**VPL AL KIFL *GPGG*	68	-0.44
EmrD	EmrD, TMH2 (27)	*GGPG* VMGAY LLTYG VSQLF YGPI *GPGG*	9	1.3
EmrL	EmrD, TMH2F (44)	*GGPG* **R**DLNV **R**EGAV QS VMGAY LLTYG VSQLF YGPI SD**R**VG *GPGG*	1	2.67

[a] TMHn denotes the nth TM helix in the protein, and TMHnF denotes the nth TMH with flanking loops.

[b] GGPG and GPGG shown in intalic are insulating flanks, and positively charged residues are highlighted in bold. The underlined segments were predicted to be TMH and restrained to be in helical conformation during the simulations.

[c] The insertion efficiencies were measured using a Sec61 translocon-based membrane integration essay by Hedin et al. [14], and ΔG_{app} was converted directly from the insertion efficiency.

Figure Legends

FIGURE 1. The PMFs of insertion into DPPC bilayers for two mTMHs with and without the flanking loops. The PMFS were calculated based on WHAM analysis of the last 40 ns of the 100-ns umbrella sampling trajectories. The z-axis represents the distance of the center of mass of the peptide is from the center of the DPPC bilayer. The thickness of the DPPC bilayer in the MARTINI force field is ~ 2 nm.

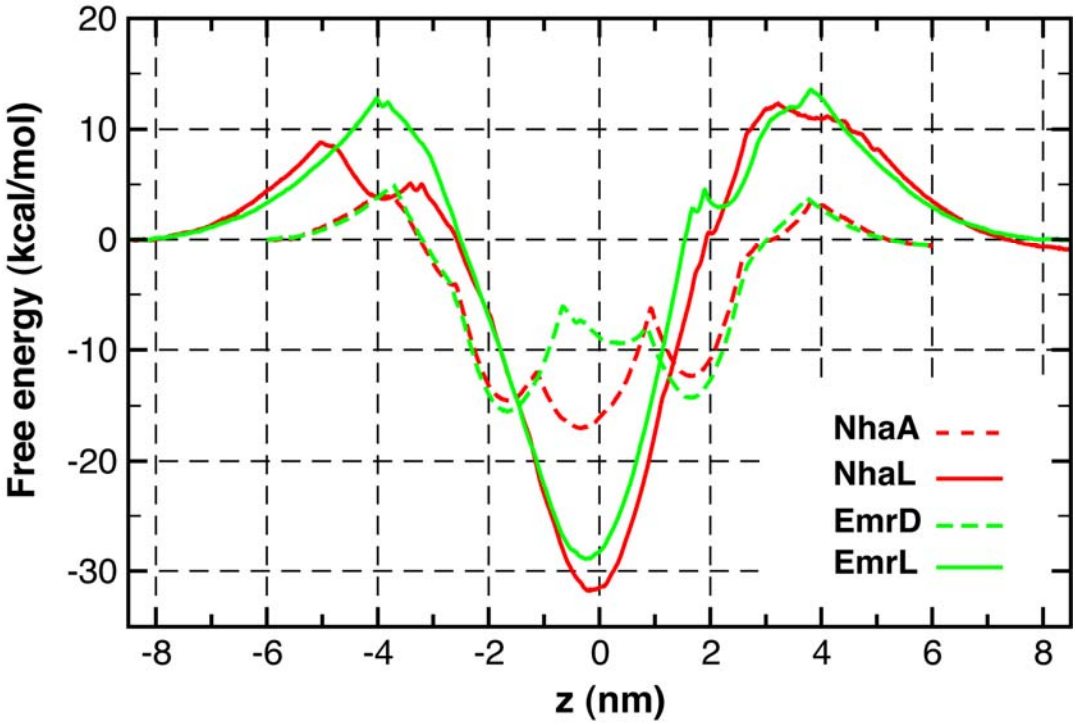


FIGURE 2. Representative snapshots of NhaA (top row) and EmrD (bottom row) at various positions of insertion into model DPPC bilayers. The peptides are shown in backbone traces, with the helical regions colored in cyan, the N-terminal loop (GGPG) in purple and the C-terminal loop (GPGG) in yellow. Negatively and positively charged side chains are also shown in blue and red spheres, respectively. For the DPPC bilayer, only the phosphate groups are shown as orange spheres for clarity. The (bulk) CG water molecules are shown as green dots, while water molecules penetrating into the bilayer and within 6.0 Å from the peptide are shown as green spheres.

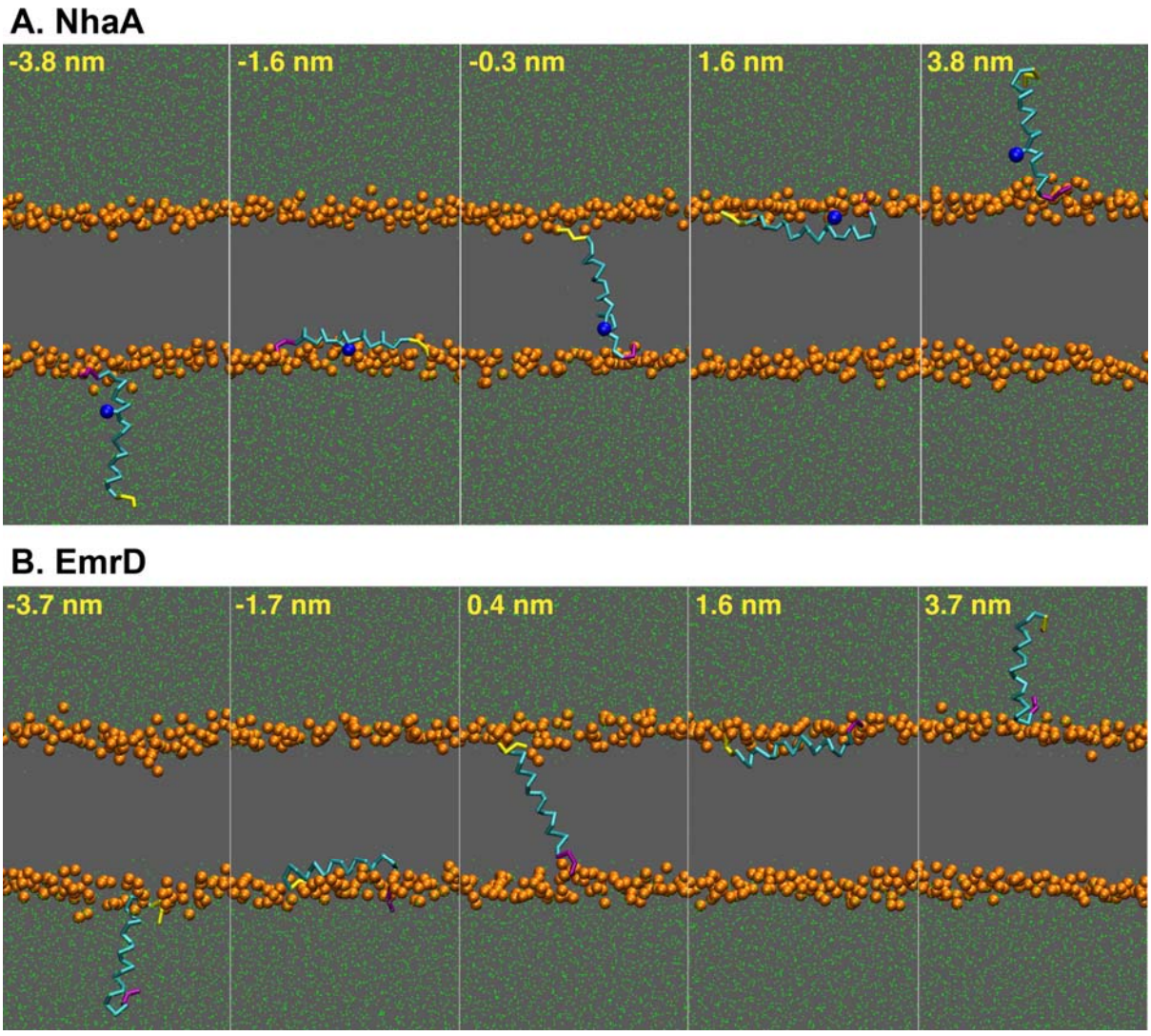


FIGURE 3. Representative snapshots of NhaL (top row) and EmrL (bottom row) at various positions of insertion into model DPPC bilayers. The same molecular representation is used as described in Fig. 2. Note the substantial compaction of EmrL and significant number of water molecules that accompany the peptides during various phases of membrane insertion.

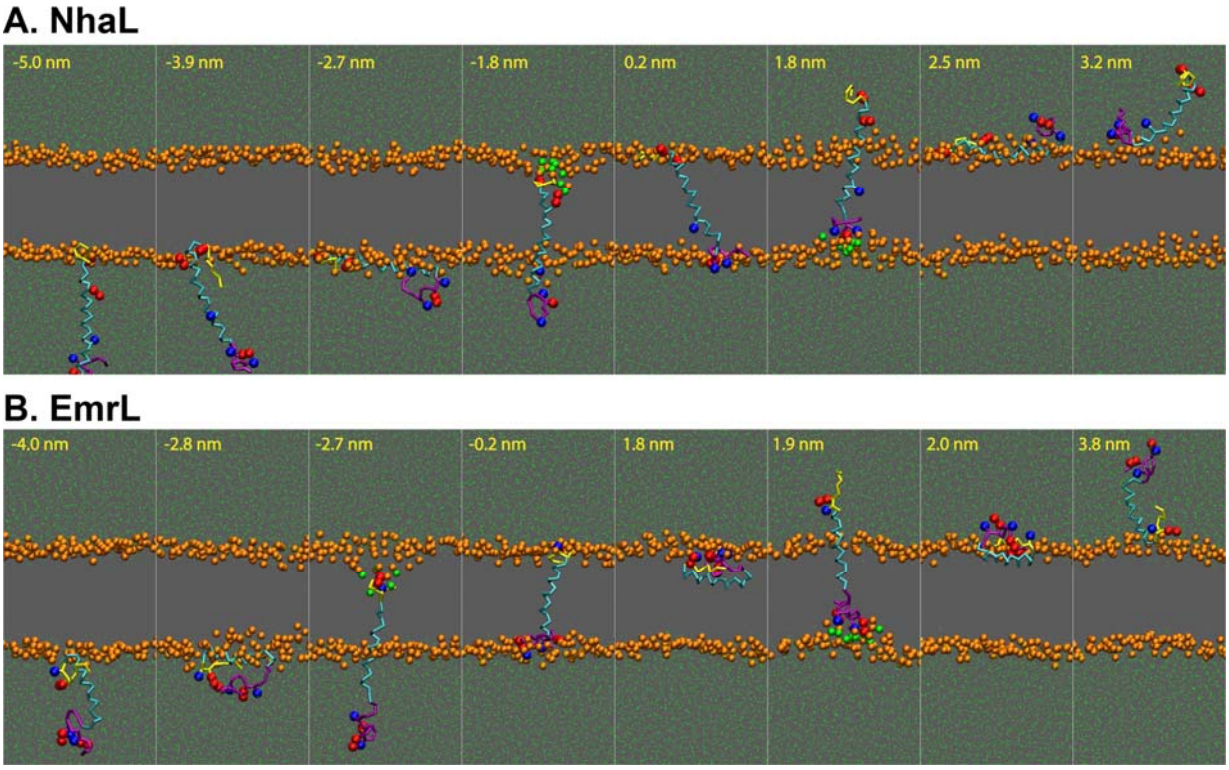


FIGURE 4. Probability distributions of the TMH tilt angle with respect to the bilayer normal as a function of the distance from the bilayer center. The distributions were calculated using the last 40 ns of the umbrella sampling trajectories.

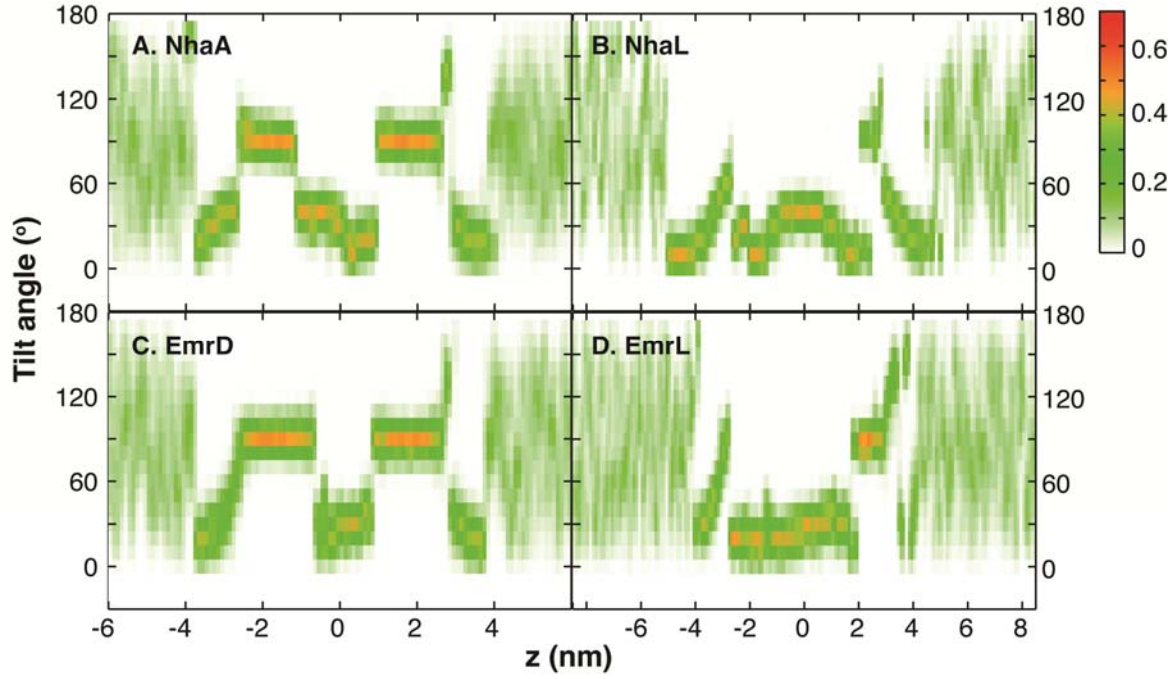


FIGURE 5. Average R_g as a function of the distance from the bilayer center for A) NhaA and NhaL, and B) EmrD and EmrL, calculated from the last 40 ns of the umbrella sampling trajectories.

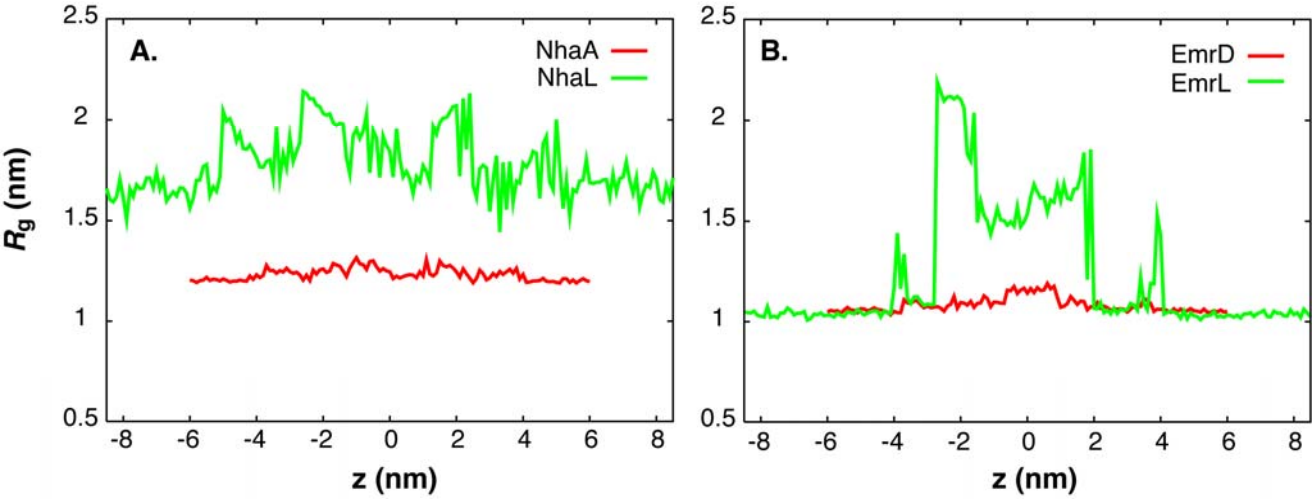


FIGURE 6. Free energy contributions arising from the interactions of the C-terminal loop (LC) and N-terminal loop (LN) with water molecules (Wat) and DPPC head (Head) and tail groups (Tail). The head group includes choline, phosphate and glycerol groups, and the rest of the DPPC molecule is defined as the tail. The results were derived from force decomposition analysis using the last 40 ns of the umbrella sampling trajectories. The cyan stripes indicate the approximate locations of the membrane/water interface.

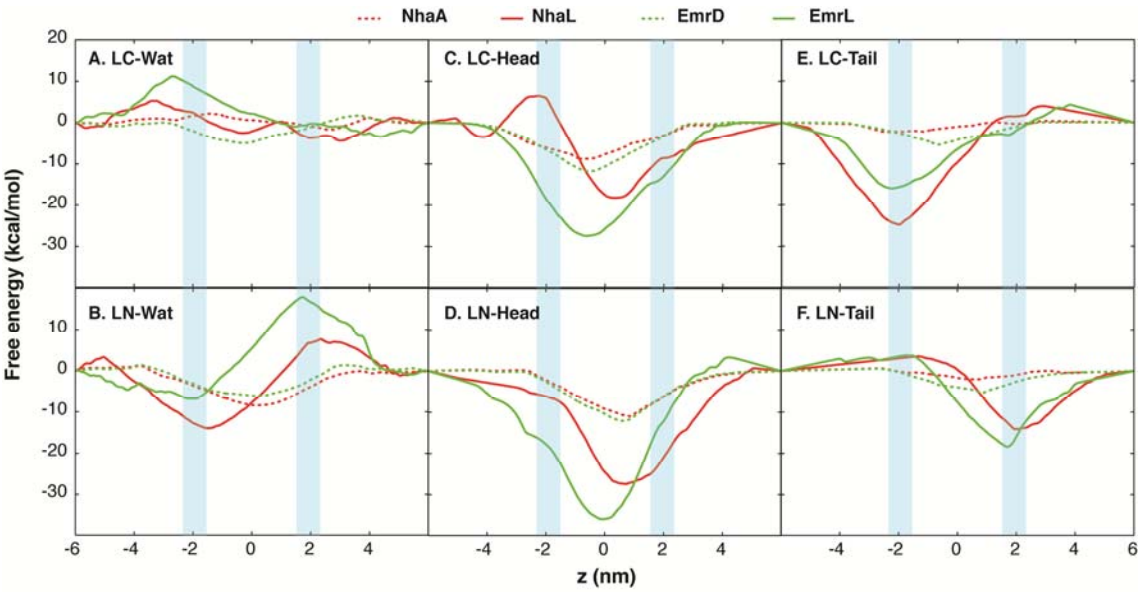


FIGURE 7. Snapshots of the final conformations of A) NhaL and B) EmrL after 30 to 50 ns simulations under the aqueous (water), interfacial (IF) and transmembrane (TM) conditions. The initial conformations (with the full simulation boxes) are shown in Fig. S3. The peptides are shown in cartoon representations with the TMH segments colored in cyan, the N-terminal loop in purple and the C-terminal loop in yellow. Negatively and positively charged side chains are also shown in blue and red sticks, respectively. Only the phosphorus atoms of the DPPC lipids are shown, as orange spheres, for clarity. The water oxygen atoms are shown as green dots.

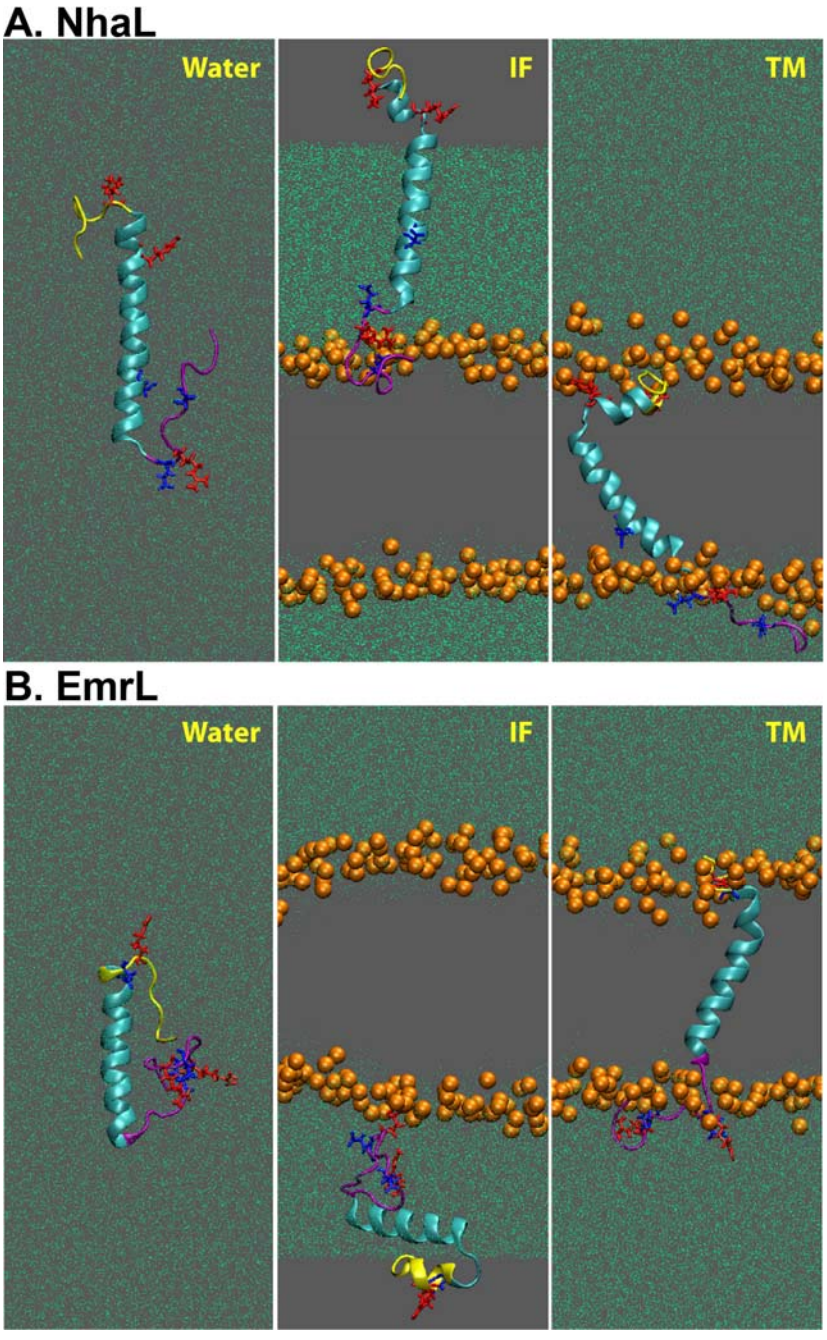
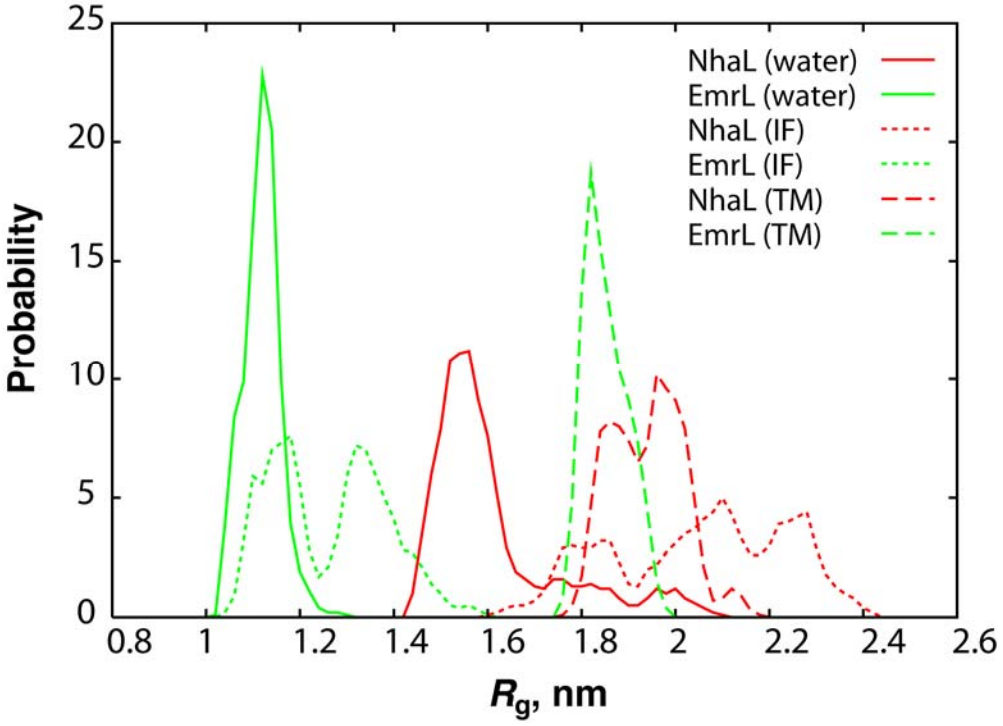


FIGURE 8. Probability distributions of R_g of NhaL and EmrL calculated from atomistic simulations in explicit water and membrane. For simulations in water, only data from the last 30 ns of the 50-ns production simulations were included; for simulations in the interfacial (IF) and transmembrane (TM) states, data from the last 20 ns of the 30-ns production simulations were included.



TOC Figure