Thermodynamics of Peptide Insertion and Aggregation in a Lipid Bilayer

Arneh Babakhani,*,† Alemayehu A. Gorfe,† Judy E. Kim,† and J. Andrew McCammon†,‡,§

Department of Chemistry & Biochemistry, Department of Pharmacology, and Howard Hughes Medical Institute, University of California at San Diego, 9500 Gilman Drive, MC 0365 La Jolla, California 92093-0365

Received: May 28, 2008

A variety of biomolecules mediate physiological processes by inserting and reorganizing in cell membranes, and the thermodynamic forces responsible for their partitioning are of great interest. Recent experiments provided valuable data on the free energy changes associated with the transfer of individual amino acids from water to membrane. However, a complete picture of the pathways and the associated changes in energy of peptide insertion into a membrane remains elusive. To this end, computational techniques supplement the experimental data with atomic-level details and shed light on the energetics of insertion. Here, we employed the technique of umbrella sampling in a total 850 ns of all-atom molecular dynamics simulations to study the free energy profile and the pathway of insertion of a model hexapeptide consisting of a tryptophan and five leucines (WL5). The computed free energy profile of the peptide as it travels from bulk solvent through the membrane core exhibits two minima: a local minimum at the water−membrane interface or the headgroup region and a global minimum at the hydrophobic−hydrophilic interface close to the lipid glycerol region. A rather small barrier of roughly 1 kcal mol^{-1} exists at the membrane core, which is explained by the enhanced flexibility of the peptide when deeply inserted. Combining our results with those in the literature, we present a thermodynamic model for peptide insertion and aggregation which involves peptide aggregation upon contact with the membrane at the solvent−lipid headgroup interface.

Introduction

The lipid bilayer plays critical roles in the biochemistry of cells, the most basic of which being its role in defining the shape of the cell and organelles. The unique arrangement of hydrophilic and hydrophobic groups at the membrane exterior and interior regions, respectively, allows for a selective inhibition of foreign agents and transport of essential molecules.[1,2] As a consequence, the membrane plays vital immunological functions.[3,4] The functional versatility of the cell membrane is also in tune with its highly inhomogeneous composition and structure, wherein receptors, channels, peptides, and other molecules create a vast chemical mixture and diverse geometry.[5] The thermodynamics of adhesion and membrane insertion of biomolecules is therefore the subject of much research. In the context of protein folding, for instance, understanding how membrane proteins insert and simultaneously fold to assume their functional three-dimensional structure is of great interest.[6] This poorly understood process can involve the cooperative action of several amino acid residues that work in concert toward both membrane penetration and organization to a particular shape.

Researchers are also studying how smaller biomolecules insert into the membrane and aggregate, and to what extent these processes affect the physical properties of the membrane.[7,8] Several experimental methods have been employed to describe the equilibrium properties of peptide insertion, that is, to describe the initial and final states and the associated thermodynamic properties.[9–11] The most important thermodynamic parameter in this context is the free energy change, ΔG, defined here to be the energetic change upon the transfer of the peptide from its initial aqueous environment to its final state in membrane; $\Delta G < 0$ indicates a favorable insertion process. Wimley et al. studied the free energy of amino acid partitioning first in an octanol/water[12] membrane mimic and later in a synthetic membrane.[13] In the latter study, they derived a thermodynamic scale for the partitioning of small model peptides between membrane and water. By altering a single residue in the test peptide, the contribution of each of the 20 amino acids to the transfer free energy was tabulated. For instance, in the case of the hydrophobic amino acid leucine, the free energy of transfer is -0.56 ± 0.04 kcal mol^{-1}. Wimley et al. also contemplated how peptide aggregation or formation of higher order molecular assemblies might occur in membranes.[14–16]

Concurrently, computational methods have played significant roles in elucidating the atomic-level details that are not always accessible by experiment. The advance of computational power and the development of faster algorithms enables modeling of increasingly larger membrane systems.[17–20] For example, Aliste and Tieleman used MD simulations to study the partitioning of the Wimley and White peptides in a model membrane, and the results provided the details of the peptide−membrane atomic interactions responsible for insertion.[21] The partitioning of the various amino acids into a model membrane host has also been studied.[22] A unique benefit of computation is its ability to shed light on the free energy of peptide−membrane binding, as well as to characterize the energetic landscape or the pathways of insertion and aggregation. One such study examined the free energy profile (or the potential of mean force, PMF) of the chromophore indole ring as it traveled through the membrane.[23] In this work, the authors explored how different parametrizations of the indole moiety can affect the free energy calculations; they also pinpointed some key properties of the reaction coordinate, including the locations of energetic barriers.

* Corresponding author. E-mail: ababakha@mccammon.ucsd.edu. Phone: (619)895-6540. Fax: (858)534-4974.
† Department of Chemistry and Biochemistry.
‡ Department of Pharmacology.
§ Howard Hughes Medical Institute.

10.1021/jp804710v CCC: $40.75
Published on Web 08/06/2008

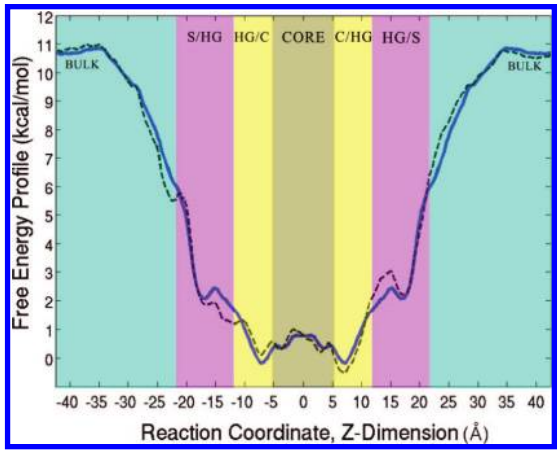

Figure 1. Free energy profile (symmetric-heavy blue line, asymmetric-dashed line) of the peptide as a function of position along the z dimension (negative/positive values correspond to the lower/upper leaflets, respectively). Aqua green on either side of the simulation box denotes regions of bulk solvent. Pink marks the regions of the solvent/lipid headgroup interface (S/HG in the lower leaflet, HG/S in the upper). Beige indicates the headgroup/core interface (HG/C in the lower leaflet, C/HG in the upper). The light brown region centered on the zero depicts the membrane core. See the text for further discussion and the computed changes in free energy.

In this study, we employ a molecular dynamics (MD) technique known as "umbrella sampling"[24,25] to simulate the insertion of a model hydrophobic peptide into a membrane. The model of choice is the small hexapeptide consisting of one tryptophan (TRP) and five leucine residues (WL5), the same model system used in the experiments of Wimley and White.[13] We computed the PMF as the peptide moves from the solvent and inserts into and completely traverses across the membrane. To our knowledge, this is the first attempt in computing the free energy profile of a full length peptide as it crosses a membrane. We analyzed the dynamical and structural properties of the peptide during this process. By coupling the computed free energy profile with the experimental results of Wimley and White, we propose a thermodynamic model for the insertion and aggregation of hydrophobic peptides in a membrane host.

Results and Discussion

Each leaflet of the membrane can be divided into three regions based on the polarity of the constituent atoms: the solvent−headgroup interface (S/HG), where the first two solvent shells ($\sim$ 6 Å) merge with the choline and phosphate head groups of the phospholipids; the glycerol or headgroup−core interface (HG/C), where the lipid head groups mix with the hydrophobic fatty acid chains; and the core, the region occupied by the aliphatic lipid tails. By using these three regions as landmarks, the free energy profile of the peptide WL5 as it traverses across the membrane is shown in Figure 1. See Supporting Information for a figure (S1) demonstrating the extent of convergence of this profile (in the allotted sampling time). The profile exhibits minima at both the S/HG and the HG/C interfaces of each leaflet, the latter being the global minimum. Note that two peptide orientations (one per leaflet) were used to determine whether the initial orientation affects the energetics of insertion. The dashed line in Figure 1 clearly shows that the initial orientation does indeed affect the profile. Such an asymmetric profile implies that even for a rudimentary peptide such as WL5, there may be more than one possible path of membrane insertion.

The conformation of the peptide, as judged by the distribution of Φ/Ψ angles, assumes that of a β strand, especially when the peptide is positioned somewhere in the membrane; see Figure S2 of Supporting Information. Most noticeably, the distribution centers on the β strand Φ/Ψ angles when the peptide is embedded in the HG/C region of the membrane (lower left panel of Figure S2). Such a conformational change may have some effect on the free energy profile of the peptide as it travels through the membrane. Yet, it is difficult to quantify this effect because we have not imposed any constraints on the Φ/Ψ angles of the peptide.

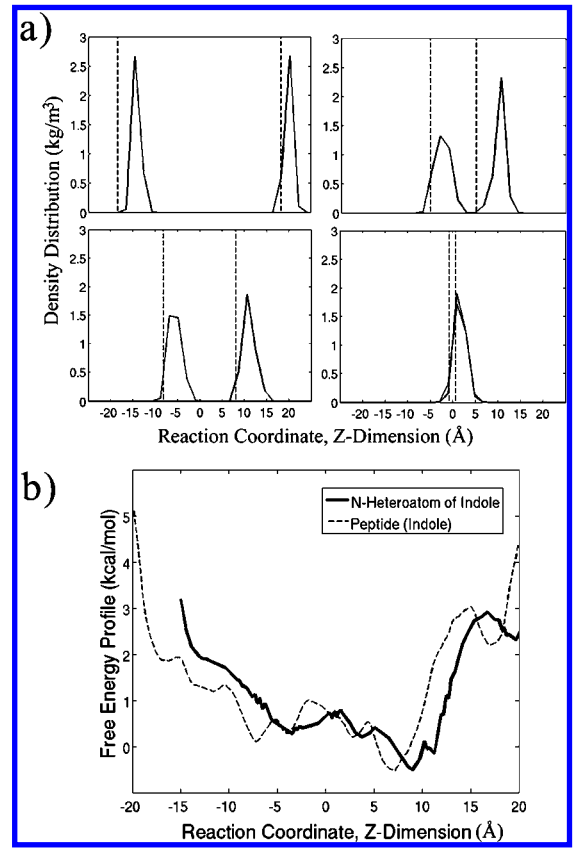

Figure 2. Nitrogen heteroatom distribution along the reaction coordinate. (a) Density distributions of the tryptophan indole nitrogen in the upper and lower leaflets of four simulation windows. Dashed lines indicate where the peptide was constrained (via the indole ring) in that particular simulation. (b) The asymmetric free energy profile in the membrane (dashed line) and the same profile shifted to reflect the nitrogen heteroatom position (dark heavy line). At each point, the nitrogen is distributed to the right (more positive z) of the indole center; but the magnitude of that shift is not equal in all parts of the profile.

In the upper leaflet, the nitrogen heteroatom points toward the interfacial regions while the peptide moves through the membrane; while in the lower leaflet, the ring nitrogen consistently points toward the core of the membrane (see Figure 2 and Figure 3). This preferential orientation results in distinct and deeper energy minima at the interfacial regions of the upper leaflet while the corresponding minima in the lower leaflet are shallow and less distinct (see Figure 1 and Figure 2b). The location of the minima and the orientation of the indole ring in the upper leaflet indicate that the ring is in a prime position to interact with the glycerol carbonyls and other hydrogen bonding elements of the lipid head groups. The importance and extent of hydrogen bonding between the indole ring and the membrane interfacial region has been addressed in a previous study.[26] Here, the nitrogen heteroatom forms a significant number of hydrogen bonds with the membrane, and it does so to a greater extent where the minima occur along the reaction coordinate (see

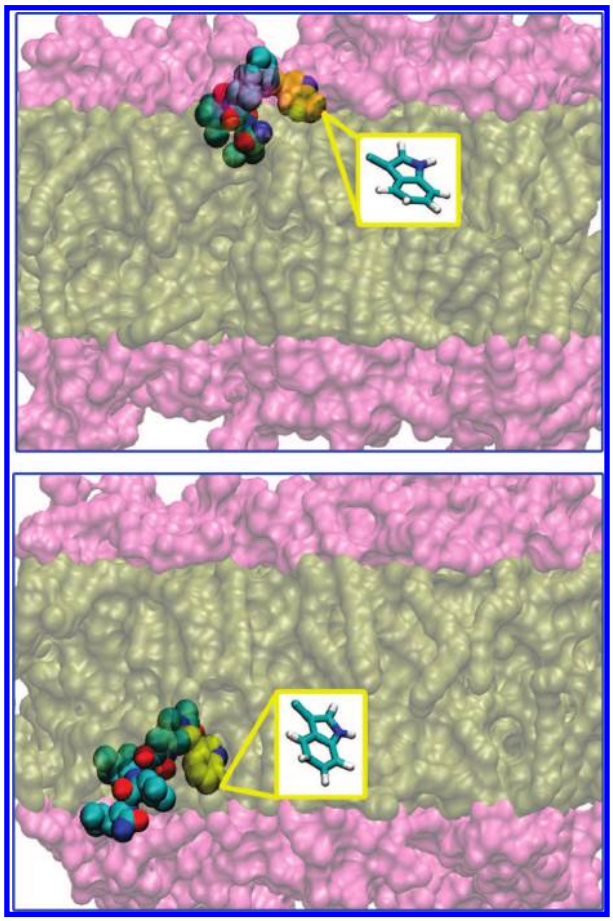

Figure 3. Snapshots of WL5 in the upper (above) and lower (below) leaflets of the membrane. The membrane is shown as a transparent surface with the lipid head groups colored pink and the tails brown. The peptide is shown as van der Waals spheres with the indole ring highlighted in yellow. The insets show a close-up of the indole ring orientation, with carbon, hydrogen, and nitrogen atoms colored in cyan, white, and blue, respectively.

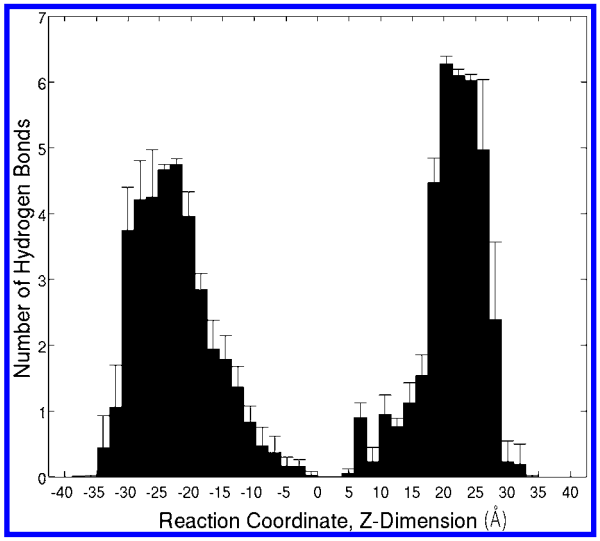

Figure 4. Average number of hydrogen bonds formed between the indole nitrogen and the hydrogen bonding elements of the lipids, as a function of the peptide's position along the reaction coordinate. The cutoff criterion was set to elements within 3 Å and hydrogen bonding angles of 40°.

Figure 4). In the upper leaflet, where the minima are deeper and more distinct, a larger number of hydrogen bonds are formed as compared with the minima in the lower leaflet.

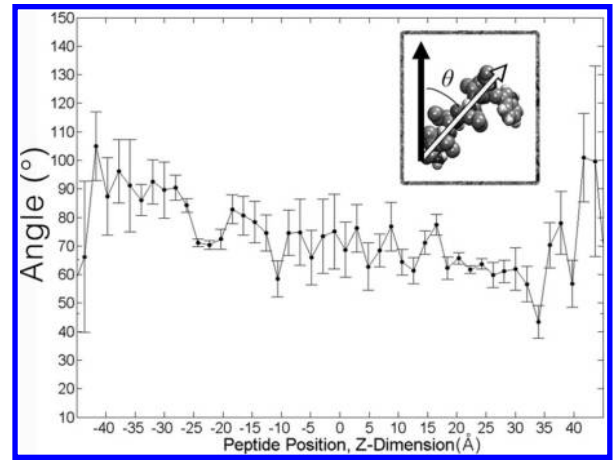

Figure 5. Peptide tilt as a function of position along the z axis, with the membrane centered on 0. Each data point corresponds to one window. Error bars encompass one standard deviation. Inset shows how the angle is defined by the vector fitted to contain the alpha carbons of the peptide (light arrow) and the vector representing the membrane normal (dark arrow).

Analysis with respect to orientation of the entire peptide reveals a tilt in the peptide axis (see Figure 5), an observation common in helical transmembrane−peptide systems.[27] As it adheres and inserts into the membrane, the peptide axis stabilizes at an angle near 70°, which is off parallel to the membrane plane (an angle of 90°). Such a tilt allows the leucine residues to interact with the hydrophobic part of the membrane while maintaining the indole ring position in the interfacial region. This appears to be the general strategy of transmembrane peptides, in that they are designed to exploit both the hydrophilic and the hydrophobic parts of the membrane during the insertion process. The orientations of the whole peptide and the tryptophan side chain argue that the PMF in the upper leaflet represents a more plausible path of peptide insertion. But in reality, the peptide could approach the membrane in a variety of orientations. Also, it may be the case that the peptide binds the membrane surface in a random initial orientation and then rearranges prior to further insertion and stabilization.

However, we do not have any a priori knowledge of the peptide orientation or information on its conformational dynamics at the solvent−membrane interface. Furthermore, the constraint applied at the TRP residue during the umbrella sampling does not allow for a spontaneous reorganization. It is therefore impossible to determine the relative weights of the two profiles, and here we assume that the lower- and upper-leaflet orientations are equally probable. In other words, because the molecular dynamics simulations in each leaflet are separate and independent, one can consider the data from the upper and lower leaflets as twice the sampling of the same system. Thus, we symmetrize the free energy profile (heavy blue line in Figure 1) and allow this result to reflect the sampling of the entire system (see Materials and Methods). This symmetric profile is used for all further calculations and discussion, but we note that this symmetry is due purely to the homogeneity of the model membrane used in the simulations. In biological heterogeneous membranes, however, such symmetry may not exist.

The peptide begins its approach toward the membrane from a region of bulk solvent, where the PMF plateaus (±35 Å). This plateau indicates that the peptide does not interact with the membrane and instead experiences a purely aqueous environment. The average PMF in this region (aqua-colored in Figure 1) is 10.1 kcal mol^{-1}. This value will be used as the

reference point from which the change in free energy (ΔG) will be calculated. In other words, this bulk solvent environment is considered as the initial state of the peptide. As the peptide travels closer to the membrane, at about 30 Å from the center, the free energy rapidly declines as the peptide encounters the solvent−headgroup (S/HG) interface. This steep descent can be attributed to a hydrophobic effect. Such a drop in the free energy is often encountered when a hydrophobic solute encounters a like environment.[28] Inside this S/HG region, the free energy profile continues to decline until it reaches a shallow minimum near ±17 Å, and the barrier to escape this minimum is approximately 0.5 kcal mol^{-1}. [This barrier seems insignificant given the sampling error of this method (see Materials and Methods), and it should be greater. Nonetheless, we know from previous experiment and simulation that a minimum in the free energy profile occurs here.] The computed free energy change as the peptide inserts into the S/HG interface is $\Delta G_{S/HG} = -8.3$ kcal mol^{-1}. Thus, the peptide has a tendency to localize in the S/HG region, and this can likely be attributed to the partial hydrophilic nature of the TRP residue.

The peptide descends further into the headgroup−core (HG/C) interface and encounters a global minimum at about ±7 Å and gains a further 2 kcal mol^{-1} in free energy. The total free energy change from the bulk solvent to the HG/C region ($\Delta G_{HG/C}$) is −10.2 kcal mol^{-1}. A local maximum occurs in the core of the membrane (0 Å), and the energy barrier as measured from the HG/C region is approximately 1.0 kcal mol^{-1}. The peptide exhibits considerable conformational changes as it travels through the membrane. The root-mean-square fluctuation (RMSF) demonstrates how much the peptide deviates from its average conformation and is plotted in Figure 6a as a function of the peptide location along the membrane normal. One can conclude that the peptide is more flexible and samples a wider region of conformational space (has a larger RMSF) as it travels deeper into the membrane core. This flexibility is also demonstrated by the root-mean-square deviation (rmsd) of the peptide structure, as seen in Figure S3 of Supporting Information. An alternative view of the RMSF in Figure 6b further shows that the side chains of the peptide are predominantly responsible for this flexibility. In particular, the TRP side chain has the largest RMSF when deep in the membrane core. To the extent that peptide flexibility can be correlated with entropy,[29] one can infer that the peptide exhibits increased entropy as it inserts into the membrane. This results in a more negative (or favorable) contribution to the free energy change and explains why the energy barrier in the core is not greater in the calculated profile. Therefore, a small barrier seems to suggest that the peptide can occasionally translate across the membrane center and shift positions between the two leaflets. Yet, as Wimley and White have demonstrated and as we shall explore further below, individual units of this model peptide aggregate in the membrane and form a supramolecular structure, a process that can counteract the tendency of a lone peptide to transfer from leaflet to leaflet.[14,15]

The free energy profile presented here is that of a single peptide traveling through the membrane and is thus not reflective of the entire biological process. If the final state of WL5 in the membrane is not monomeric but rather an aggregate form, then it behooves us to study this aggregation process and the pathways that lead to it. Although we have not simulated the actual aggregation, the presented work, coupled with previous simulations and experimental evidence, can shed some light on the mechanism and potential pathways of aggregation. Figure 7 shows a proposed thermodynamic cycle for the insertion and

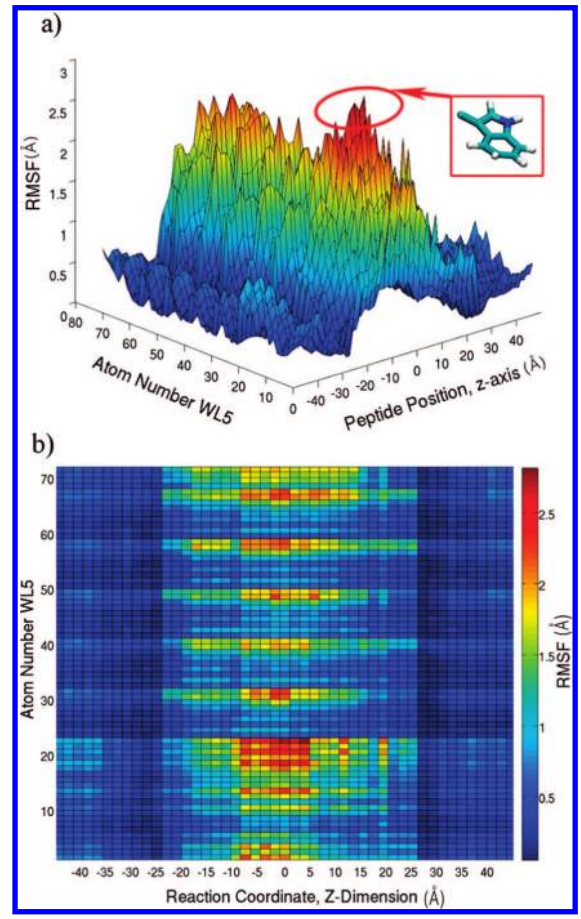

Figure 6. Root mean square fluctuation (RMSF) of the peptide by atom number and as a function of position along the z axis of the simulation box. The membrane is centered on the zero mark of the z axis. Atom numbers 1−3 correspond to the acetyl group on the N-terminus; 4−24 correspond to the tryptophan residue; 25−69 correspond to the leucine residues; and 70−72 correspond to the amide group on the C-terminus. (a) Surface representation, where the red halo shows the maximum RMSF achieved by the indole ring near the core of the membrane. (b) An alternative (color map) view of the RMSF data.

assembly of WL5 in the membrane. In this model, the initial state of the peptide in an aqueous environment is monomeric, and the final state in membrane is an assembly of several monomers (top left and bottom right of Figure 7, respectively). The difference in the free energy between these two states as obtained by Wimley and White is $\Delta G_{Exp} = -5.3$ kcal mol^{-1} (Exp = experimental).[13] [This is assuming that the transfer from monomeric aqueous to membrane aggregate form is an equilibrium process, and that the experimental result is a reflection of this.] As suggested by Grossfield et al., the computed free energy can be coupled to this experimental value via a correction term that accounts for the difference in the peptide and lipid concentrations between simulation and experiment.[30]

The correction term is given by

$$\Delta G_{Corr} = \Delta G_R + \Delta G_{Mix} \quad (1)$$

$$\Delta G_R = k_b T \ln(\chi_{sim,Peptide}/\chi_{exp,Peptide}) \quad (2)$$

$$\Delta G_{Mix} = k_b T \ln(\chi_{sim,lipids}/\chi_{exp,lipids}) \quad (3)$$

where k_bT is the Boltzmann factor. Following the notation of Grossfield, ΔG_R and ΔG_{Mix} are the correction terms for the peptide and lipid (respectively) concentration differences between simulation (sim) and experiment (exp). χ is the mole

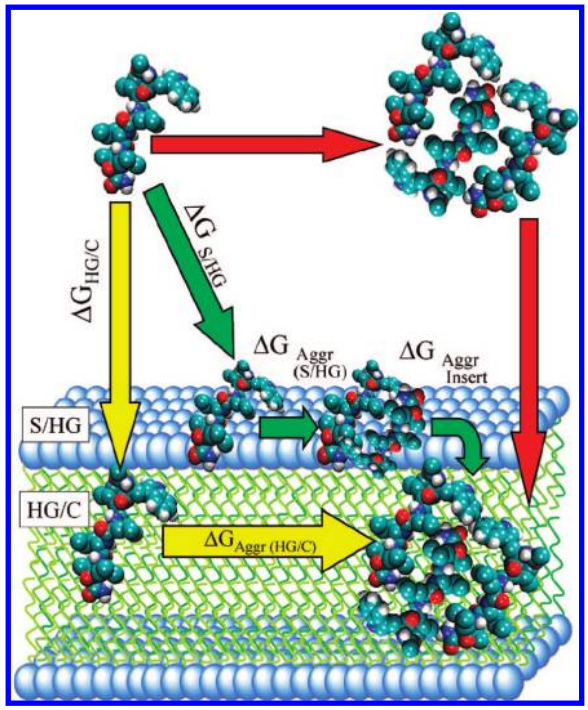

Figure 7. Proposed thermodynamic cycle of peptide insertion and supramolecular assembly inside a model membrane. The yellow arrows depict a pathway in which single peptides insert first into the deep headgroup−core (HG/C) interface and then aggregate to form the final structure. The red arrows mark a less likely (yet plausible) path, in which the peptide aggregates first in solution, and then the aggregate as a whole inserts into the membrane. The green arrows qualitatively describe an intermediary pathway, in which individual peptides insert first into the aqueous or solvent−headgroup (S/HG) interface, aggregate to some extent, and then insert deeper into the membrane. Each of these steps can be described by a free energy change (ΔG, Aggr = aggregation). See the text for further discussion.

fraction of the species in question (either peptide or lipids). This entire correction term, ΔG_{Corr}, is calculated to be +3.0 kcal mol^{-1} (see Materials and Methods for further information) and must be added to the computational results.

In terms of insertion and aggregation, Figure 7 shows two paths on opposite ends of the spectrum. In the yellow path, each peptide inserts individually into the deeper membrane interface. The aggregation step then follows:

$$\Delta G_{\text{HG/C}} + \Delta G_{\text{Aggr(HG/C)}} = \Delta G_{\text{YellowPath}} \quad (4)$$

$$(\Delta G_{\text{HG/C}} + \Delta G_{\text{Corr}}) + \Delta G_{\text{Aggr(HG/C)}} = \Delta G_{\text{YellowPath}} \quad (5)$$

Equation 5 is the corrected version of eq 4, where the correction term has been added to computational result $\Delta G_{\text{HG/C}} = -10.2$ kcal mol^{-1}. The unknown term here is the aggregation one, $\Delta G_{\text{Aggr(HG/C)}}$. If we can set $\Delta G_{\text{YellowPath}}$ to be the experimental value determined by Wimley and White (−5.3 kcal mol^{-1}), the aggregation term is then calculated to be +1.9 kcal mol^{-1}.

The red path in Figure 7 illustrates the opposite approach, in which individual peptides in aqueous solution first assemble, and then the whole aggregate inserts into the membrane. The first leg of this pathway is plausible and has a negative free energy change, because hydrophobic peptides in an aqueous environment are energetically more stable in aggregate instead of monomeric form.[31] Yet, the stability and energetic gain of the aggregate in solution could deter the adhesion and insertion of the aggregate into the membrane, thus making the second leg of the red path less likely. Furthermore, if the orientation of each monomeric peptide can affect the thermodynamics of the insertion process (as discussed above), it may become energetically costly for the peptide units of the complex to reorganize and assume the correct orientation. Others have also shown that transmembrane proteins do not fold entirely or form complex structures before inserting into a membrane.[32]

Instead, it is more likely that a transmembrane peptide adsorbs on the membrane surface, aggregates to some extent, and then inserts into the membrane. As the insertion event proceeds, it is likely that the membrane continues to shape the peptide aggregate structure until it achieves a final stable form in the membrane host. Note that membranes exhibit the mechanical properties to do just this.[33] For our model system, this is depicted in the green path, in which the monomeric peptide first adheres to the S/HG interface, aggregates, and then inserts:

$$\Delta G_{\text{S/HG}} + \Delta G_{\text{Aggr(S/HG)}} + \Delta G_{\text{Aggr(Insert)}} = \Delta G_{\text{GreenPath}} \quad (6)$$

$$(\Delta G_{\text{S/HG}} + \Delta G_{\text{Corr}}) + \Delta G_{\text{Aggr(S/HG)}} + \Delta G_{\text{Aggr(Insert)}} = \Delta G_{\text{GreenPath}} \quad (7)$$

Equation 7 represents the corrected version, with the correction term added again to the computational result, $\Delta G_{\text{S/HG}} = -8.3$ kcal mol^{-1}. The last two terms on the left-hand side of equation 7 are the unknowns, and we cannot calculate the exact contribution from $\Delta G_{\text{Aggr(S/HG)}}$ or $\Delta G_{\text{Aggr(Insert)}}$. Again, if we equate $\Delta G_{\text{GreenPath}}$ to the experimental value of −5.3 kcal mol^{-1}, the sum of these unknown terms can be calculated, $\Delta G_{\text{Aggr(S/HG)}} + \Delta G_{\text{Aggr(Insert)}} = 0$ kcal mol^{-1}.

Thus, in the yellow path, the aggregation free energy change is positive and indicative of an unfavorable process. In other words, if the monomeric form of the peptide were to insert deep into the HG/C region of the membrane, aggregation would not be likely. In the green path however, the collective aggregation term is 0, indicating an equilibrium (and more likely) process. We have shown the following relation:

$$\{\Delta G_{\text{Aggr(S/HG)}} + \Delta G_{\text{Aggr(Insert)}}\}_{\text{GreenPath}} < \{\Delta G_{\text{Aggr(HG/C)}}\}_{\text{YellowPath}} \quad (8)$$

where the braces group the unknown terms of each path, as described above. The results suggest that the green path in Figure 7, where the peptide first adheres to the solvent−headgroup interface, aggregates, and then inserts, more closely represents the in vitro mechanism of insertion and organization of WL5. Furthermore, following the previous discussion of the red path in which we stated that a formed aggregate would have difficulty inserting into the membrane, we can speculate that $\Delta G_{\text{Aggr(Insert)}} > 0$ and thus $\Delta G_{\text{Aggr(S/HG)}} < 0$. In other words, aggregation occurs immediately upon peptide binding to the membrane, in the solvent−headgroup region.

Materials and Methods

System Construction. The structure of the solvated peptide−membrane system was obtained from a previous study, in which we explored the properties of the system in a 75 ns MD simulation.[25] Snapshots at the end of this simulation, at which point, the peptides were embedded within the headgroup region of the membrane, were used as starting configurations to initiate 50 different all-atom simulations each representing a window in an umbrella sampling scheme. Two different snapshots of the peptide were employed (one for each leaflet of the membrane); in other words, the 25 simulation windows within a leaflet were initiated with the same initial peptide conformation. These two snapshots are taken to be the equilibrated

conformation of the peptide when embedded in the solvent−headgroup interfacial region of the membrane.

In total then, the 50 simulation windows consisted of a WL5 peptide (acetylated and amidated at its N- and C-termini, respectively), a preconstructed bilayer of 128 dimyristoylphosphatidylcholine (DMPC) lipids,[20] and 6631 water molecules, resulting in a ∼25 000 atom system assembled in a simulation box of 60 × 60 × 95 Å. The membrane was positioned at the center of the box, leaving approximately 23 Å for the solvent on either side.

Each simulation corresponds to a window of width 1.9 Å such that the position of WL5 was different in each simulation. In the first window, WL5 was constrained (along the z axis) in the bulk solvent at the upper half of the box and well away from the membrane surface. In the next, WL5 was moved by 1.9 Å closer to the membrane and constrained at that position. In subsequent simulations, WL5 was moved closer to the center of the simulation box each time stepped along the z axis by the increment of 1.9 Å. The constrained molecule thus spans the entire reaction coordinate (in the z direction) of approximately 95 Å.

For organizational purposes, the simulation box is divided into two halves along the z axis, representing the "upper" (positive z) and "lower" (negative z) leaflets with respect to the membrane center at $z = 0$. Among the 25 simulation windows in each leaflet, 12 had the peptide constrained in solvent and 13 were carried out with the peptide constrained at various z locations in the membrane. The initial configuration of the peptide differs between the two leaflets but is the same among the 25 simulation windows of each leaflet. The reaction coordinate was defined by the separation along the z axis between the centers-of-mass of the indole chromophore, the tryptophan side chain, and the membrane. WL5 was constrained by a harmonic potential of force constant 500 kJ mol^{-1} nm^{-2} (1.20 kcal mol^{-1} Å^{-2}).

Simulation Setup. The setup of the simulations was as follows. First, unfavorable contacts were relieved by two cycles of 5000 steps steepest descent followed by 5000 steps of conjugate-gradient energy minimizations, each with the peptide held fixed and set free. Second, except when the peptide is in bulk solvent, an annealing step was conducted to equilibrate the lipid tails around the peptide. With the peptide and lipid head groups restrained, the system was gradually heated (over 1 ns) from 310 K (the temperature of the initial snapshot) to 410 K, then gradually cooled back down to 310 K. Third, adding pressure coupling, a restrained MD was performed on all 50 windows for 1 ns, with the restraint applied to nonsolvent molecules. Fourth, with only the indole ring of tryptophan constrained to its selected position, production runs commenced. Those with the peptide located outside of the membrane were sampled for 5 ns, while those with the peptide somewhere in the membrane were sampled for 25 ns. In total, including all pre- and post production runs, approximately 850 ns of all-atom MD simulations have been carried out.

A time step of 2 fs was used while coordinates/velocities were recorded every 500 steps (1 ps). Constraints were imposed using the LINCS method. Full electrostatics were calculated using the particle mesh Ewald (PME) method, with Coulombic and van der Waals cut-offs of 0.9 and 1.4 nm, respectively. A sixth-order spline interpolation was used for PME along with a tolerance of 1×10^{-5}. Nearest neighbor lists were updated every 10 steps using the grid method and periodic boundary conditions (in xyz) were employed with a cutoff of 0.9 nm. Berendsen temperature coupling was used, with a reference temperature of 310 K and a coupling of 0.1 ps. A semi-isotropic pressure coupling scheme, in which the $x-y$ dimensions are coupled together while the z direction is allowed to fluctuate independently, was used with a reference pressure of 1.0 bar, a pressure coupling of 0.5 ps, and a compressibility of 4.5×10^{-5} bar^{-1}.

The simulations were performed with the GROMACS MD package.[34] The ffgmx force field was used for the peptide and the SPC water model was used as the solvent. The lipid parameters of Berger et al. were employed for the DMPC lipids.[35] The analysis of the simulation trajectories was completed using the various GROMACS tools. For further details regarding these tools and specific procedures in running GROMACS MD simulations; see the GROMACS manual at http://www.gromacs.org/. Visualization and rendering were done by the Visual Molecular Dynamics (VMD) application.[36]

Symmetrization of the Free Energy Profile in Figure 1. Given the construction of the simulation windows, one might expect the resulting profile in Figure 1 to be perfectly symmetric about the membrane center. In other words, the profile in one leaflet should be a mirror image of the other provided that the peptide exhibits the same orientation while penetrating through both leaflets of a homogeneous membrane. As mentioned in the text, the free energy profile of Figure 1 is not exactly symmetric, and this was attributed mainly to the starting conformation of the peptide in the simulation windows of each leaflet. Yet, even if the peptide conformation was the same in each leaflet, it is doubtful that one could achieve great symmetry, simply due to convergence limitations inherent in MD simulations. An exorbitant amount of computational time would be needed to achieve such a result. In fact, the asymmetric profile may be used as a convergence criterion with respect to certain free energy calculations.

For instance, the values of the computed $\Delta G_{HG/C}$ in the upper and lower leaflets are within 1 kcal mol^{-1} of each other while the values of $\Delta G_{S/HG}$ differ more significantly. This could signify that the simulation of deep insertion (from bulk solvent into the deep interface) produces a more converged result than insertion into the solvent−headgroup (S/HG) interface. One could reason that this is the case because the S/HG interface is more chemically diverse and dynamic, thus requiring more sampling time for energetic calculations. In Figure 1, the symmetric profile was obtained by averaging the results of the two leaflets. About the center of the profile (at 0 Å), the average was taken between the two points on either side of the profile. For instance, if the free energy was 1.8 kcal mol^{-1} at +10 Å (upper leaflet) and 1.5 kcal mol^{-1} at −10 Å (lower leaflet), the free energy in the symmetric plot at ±10 Å is taken to be 1.65 kcal mol^{-1}. The asymmetric free energy profile in Figure 1 can also be used as a judge of error in these simulations. The differences at mirror sites along the reaction coordinate (for instance, at ±7 Å) reveal a sampling error on the order of 0.5−1.0 kcal mol^{-1}.

Calculation of the Correction Term (ΔG_{Corr}). As reported by Wimley and White, the experimental concentration of the WL5 peptide is (at most) 100 μM.[13] Using this concentration, and that of water to be 55.5 M (or 55.5×10^6 μM), we can calculate the mole fraction of the peptide in the experiment to be $\chi_{exp,Peptide} = 1.80 \times 10^{-6}$. In each simulation, we have one peptide, in either the upper or lower leaflet of the system. Because of periodic boundary conditions, the peptide is solvated in the total solvent of 6631 water molecules. Thus, the mole fraction of the peptide in the simulation is $\chi_{sim,Peptide} = 1.51 \times 10^{-4}$. Substituting these mole fractions along with the simulation temperature (310 K) into the expression for ΔG_R, we arrive at

+2.73 kcal mol^{-1}, which corrects for the differences in peptide concentration between simulation and experiment. To calculate the mole fractions of the lipids, a volumetric ratio is computed in a fashion similar to Grossfield et al. Taking the average thickness of a DMPC membrane to be 36 Å[37] and the cross-sectional area to be 60 × 60 Å (per simulation setup), we compute the volume occupied by the membrane to be 1.30 × 10^5 Å^3. When the ratio of the membrane volume to the box volume is computed, we obtain $\chi_{sim,lipids} = 0.38$; we also use $\chi_{exp,lipids} = 0.24$, for an experimental vesicle concentration of 4 mM.[29] These mole fractions result in $\Delta G_{Mix} = +0.28$ kcal mol^{-1}, and this corrects for the differences in lipid concentration between simulation and experiment. The combined correction term is then the sum, $\Delta G_R + \Delta G_{Mix} = \Delta G_{Corr} \approx +3.0$ kcal mol^{-1}.

Concluding Remarks

In summary, in this paper, we have examined the biologically important process of peptide insertion into a membrane host, followed by aggregation and supramolecular assembly. The calculated PMF of the insertion process shows two minima, one each at the solvent−headgroup and hydrophilic−hydrophobic interfaces. The energetics show that aggregation and supramolecular assembly can begin immediately when the peptide makes contacts with the membrane at the first solvent−headgroup interfacial region. In other words, the peptide need not insert deeply and stabilize before taking on higher order structure; if it does insert deeply, the aggregation process is thermodynamically unfavorable. This implies that the membrane environment plays a direct and significant role in shaping transmembrane protein/peptide structures and is not a passive medium. Even initial contacts with the membrane can induce tertiary configuration in the protein sequence; this induction can be attributed to the fluidity and chemical diversity of the solvent−membrane interface.

Acknowledgment. This work was supported in part by the National Science Foundation, the National Institutes of Health (GM34921), the Howard Hughes Medical Institute, the National Biomedical Computing Resource, the Keck Foundation, and the Center for Theoretical and Biological Physics.

Supporting Information Available: Figure showing the extent of convergence of the free energy profile; figure showing the phi/psi angles of the internal leucine residues of the peptide, at various locations in the simulation box; figure showing the rmsd of the peptide along the reaction coordinate within the membrane. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Shai, Y. *Biochim. Biophys. Acta* **1999**, *1462*, 55–70.
(2) Aungst, B. J. *J. Pharm. Sci.* **2006**, *82*, 979–987.
(3) Dykstra, M.; Cherukuri, A.; Sohn, H. W.; Tzeng, S. J.; Pierce, S. K. *Annu. Rev. Immunol.* **2003**, *21*, 457–481.
(4) Rehermann, B.; Nascimbeni, M. *Nat. Rev. Immunol.* **2004**, *5*, 215–229.
(5) Mukherjee, S.; Maxfield, F. R. *Annu. Rev. Cell. Dev. Biol.* **2004**, *20*, 839–866.
(6) Stevens, T. J.; Mizuguchi, K.; Arkin, I. T. *Protein Sci.* **2004**, *13*, 3028–3037.
(7) Tang, M.; Waring, A. J.; Hong, M. *J. Am. Chem. Soc.* **2007**, *129*, 11438–11446.
(8) Bringezu, F.; Wen, S.; Dante, S.; Hauss, T.; Majerowicz, M.; Waring, A. *Biochemistry* **2007**, *46*, 5678–5686.
(9) Alaouie, A. M.; Lewis, R. N.; McElhaney, R. N. *Langmuir* **2007**, *23*, 7229–7234.
(10) Caesar, C. E.; Esbjorner, E. K.; Lincoln, R.; Norden, B. *Biochemistry* **2006**, *45*, 7682–7692.
(11) Hong, M. *J. Phys. Chem. B* **2007**, *111*, 10340–10351.
(12) Wimley, W. C.; Creamer, T. P.; White, S. H. *Biochemistry* **1996**, *35*, 5109–5124.
(13) Wimley, W. C.; White, S. H. *Nat. Struct. Biol.* **1996**, *3*, 842–848.
(14) Wimley, W. C.; Hristova, K.; Ladokhin, A. S.; Silvestro, L.; Axelsen, P. H.; White, S. H. *J. Mol. Biol.* **1998**, *277*, 1091–1110.
(15) Wimley, W. C.; White, S. H. *J. Mol. Biol.* **2004**, *342*, 703–711.
(16) Han, X.; Hristova, K.; Wimley, W. C. *Biophys. J.* **2008**, *94*, 492–505.
(17) Bond, P. J.; Sansom, M. S. *J. Am. Chem. Soc.* **2006**, *128*, 2697–2704.
(18) Ayton, G. S.; Blood, P. D.; Voth, G. A. *Biophys. J.* **2007**, *92*, 3595–3602.
(19) Gorfe, A. A.; Babakhani, A.; McCammon, J. A. *J. Am. Chem. Soc.* **2007**, *129*, 12280–12286.
(20) Gorfe, A. A.; Babakhani, A.; McCammon, J. A. *Angew. Chem., Int. Ed.* **2007**, *46*, 8234–8237.
(21) Aliste, M. P.; Tieleman, D. P. *BMC Biochem.* **2005**, *6*, 30.
(22) MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P. *J. Gen. Physiol.* **2007**, *129*, 371–377.
(23) Norman, K. E.; Nymeyer, H. *Biophys. J.* **2006**, *91*, 2046–2054.
(24) Bartels, C.; Karplus, M. *J. Comput. Chem.* **1997**, *18*, 1450–1462.
(25) Shroll, R. M.; Smith, D. E. *J. Chem. Phys.* **1999**, *110*, 8295–8302.
(26) Babakhani, A.; Gorfe, A. A.; Gullingsrud, J.; Kim, J. E.; McCammon, J. A. *Biopolymers* **2007**, *85*, 490–497.
(27) Ramamoorthy, A.; Kandasamy, S. K.; Lee, D. K.; Kidambi, S.; Larson, R. G. *Biochemistry* **2007**, *46*, 965–975.
(28) Chandler, D. *Nature* **2005**, *437*, 640–647.
(29) Li, F.; Gangal, M.; Juliano, C.; Gorfain, E.; Taylor, S. S.; Johnson, D. A. *J. Mol. Biol.* **2005**, *315*, 459–469.
(30) Grossfield, A.; Woolf, T. B. *Langmuir* **2002**, *18*, 198–210.
(31) Narita, M.; Chen, J. Y.; Sato, H.; Kim, Y. *Bull. Chem. Soc. Jpn.* **1985**, *58*, 2494–2501.
(32) Stevens, T. J.; Mizuguchi, K.; Arkin, I. T. *Protein Sci.* **2004**, *13*, 3028–3037.
(33) Gullingsrud, J.; Schulten, K. *Biophys. J.* **2004**, *86*, 3496–3509.
(34) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
(35) Berger, O.; Edholm, O.; Jähnig, F. *Biophys. J.* **1997**, *72*, 2002–2013.
(36) Humphrey, W.; Dalke, A.; Schulten, K. *J. Molec. Graphics* **1996**, *14*, 33–38.
(37) Saccani, J.; Castano, S.; Desbat, B.; Blaudez, D. *Biophys. J.* **2003**, *85*, 3781–3787.

JP804710V