Supporting Information

Jaud et al. 10.1073/pnas.0900638106

SI Text

Molecular Dynamics (MD): Analysis of Hydrophobic Mismatch Energetics. To quantify the shape of the bilayer around every peptide, we have selected a model of lipid–protein interaction based on the Landau-de Gennes theory, in which the free energy of lipid–protein interactions is determined by minimizing the functional integral of the local free energy density (1). Doing so requires solving $d_l(r)$, the lipid thickness as a function of the radial distance from the protein, *r*. We must first assume that the peptide is fixed in the membrane and is only acting to provide a boundary condition for the lipid. Under relaxed boundary conditions ($d_l(0) = d_p$), and by using a free energy density model with terms for elastic deformation and local thickness fluctuation, one finds (2):

$$d_l(r) = d_l^0 - (d_l^0 - d_p) \exp(-r/\xi_p)$$
 [S1]

where d_p is the hydrophobic length of the peptide, d_l^0 is the unperturbed lipid thickness (far from the peptide), and ζ_p is the coherence length of the peptide perturbation.

We used this model to parameterize membrane perturbation around each simulated system, as follows. We first quantified the shape of the bilayer. Taking periodic boundary conditions into account, we selected choline groups within 50 Å of the centerof-mass (c-o-m) of the helix (orange dots, Fig. 5 Left). We then calculated the average positions of 2-Å-thick rings of cholines centered on the helix c-o-m (black dots, Fig. 5 Left). Finally, we calculated the average position of Leu sidechains every 1 Å along the z-axis (red circles, Fig. 5). Using these data, we subtracted the average lower bilayer surface from average upper surface to obtain the hydrophobic thickness as a function of the radial distance (r) from the peptide c-o-m. We ignored data for r < 5Å, because the average radial position of the Leu sidechains is \approx 5 Å (Fig. 5 *Left*). Our analysis illustrates the extensive bilayer rearrangement occurring around the peptides to maximize favorable hydrophobic interactions. The bilayer collapses around peptides to form cavities that are typically deepest for the shortest peptides.

Information gained by using Eq. **S1** can then be used to calculate the difference in free energy of deformation of the lipid-protein system:

$$\Delta G = k \left(\frac{2r_p}{\xi_l} + 1\right) (d_l^0 - d_p)^2$$
[S2]

where k is a phenomenological constant related to the bilayer area compressibility modulus, r_p is the radius of the peptide, which is assumed to have a cylindrical shape, and ζ_l is the persistence length of the lipid-bilayer fluctuations. Hence, ΔG is proportional to the square of the hydrophobic mismatch between the peptide and the unperturbed lipid, and calculating it requires the knowledge of k and ζ_l , which are not readily available. After assuming that $r_p = 5$ Å, which is the average radial distance of helical side chains from the c-o-m of the helix, that k is constant in all systems, which are made of the same type of lipid, and that ζ_p can be used as a fair approximation for ζ_l , we calculated the difference in free energy of all systems relative to G20, $\Delta G_{\text{peptide}}$ / ΔG_{GL20} . In the case of GLn, this ratio is remarkably linear, with a slope of -1.8 units per residue, an intercept of 35 units (R = 0.966). After ignoring KL6, it is also linear for KLn, with a slope of -1.3 units per residue and an intercept of 27 units (R = 0.976). Unsurprisingly, the longest peptides have the lowest free-energy of bending because they do not require their bilayers to bend as much as those around the shortest peptides.

Methods: In Vitro Expression. *Enzymes and chemicals.* All enzymes, plasmid pGEM1, DTT (DTT), and the TNT SP6 Quick Coupled Transcription/Translation System Kit were from Promega. [35S]-Met, ribonucleotides, and deoxyribonucleotides were from Amersham Pharmacia. BigDye Terminator v1.1 Cycle Sequencing Kit was from Applied Biosystems. Oligonucleotides were from Cybergene AB and MWG Biotech.

DNA manipulations. For expression of H-segment-containing Lep constructs from the pGEM1 plasmid, the 5' end of the lep gene from *Escherichia coli* was modified by the introduction of an XbaI site and by changing the context of 5' to the initiator ATG codon to a Kozak consensus sequence (3). Site-directed mutagenesis was performed by using the QuikChange Site-Directed Mutagenesis Kit from Stratagene to introduce acceptor sites for N-linked glycosylation in positions 96–98 (Asn-Ser-Thr) and 258–260 (Asn-Ala-Thr).

Oligonucleotides encoding the different H-segments were introduced between a SpeI-site in codons 226–227 and the KpnI site in codon 253 in the Lep gene (4). H-segments were constructed by using 1 or 2 double-stranded oligonucleotides with overlapping overhangs at the ends. Pairs of complementary oligonucleotides were first annealed at 85 °C for 10 min followed by slow cooling to 30 °C, after which the 2 annealed doublestranded oligos were mixed, incubated at 65 °C for 5 min, cooled slowly to room temperature, and ligated into the vector. All H-segment inserts were confirmed by sequencing of plasmid DNA at BM Labbet AB.

Peptide Synthesis. The sequences of the peptides were 6, 8, 10, and 12 Leu flanked by the groups GGPG or KKPK. In addition, each peptide carried a Trp residue on the C terminal to facilitate measurements of concentration. The peptides were synthesized on a 433A Applied Biosystems automatic synthesizer with stepwise solid-phase procedures (5) by using fluorenylmethoxycarbonyl (Fmoc) chemistry and trifluoroacetic cleavage. Syntheses were performed on a 0.1-mmol scale starting with Fmoc-Rink amide 4-methylbenzhydrylamine hydrochloride resin. Fmocprotected amino acids were used for all coupling reactions. The crude peptides were purified by reverse-phase HPLC on a preparative C18 reverse-phase column with gradients of acetonitrile in 0.1%. Peptide identities were confirmed by MALDI mass spectrometry (Chemistry Department, University of California, Irvine). Fmoc amino acids and resins for peptide synthesis were obtained from NovaBiochem (EMD Biosciences). All chemicals were of analytical reagent grade.

Circular Dichroism Spectroscopy.*Materials*. Palmitoyloleoyl phosphocholine (POPC) and palmitoyloleoyl phosphatidylglycerol (POPG) were purchased from Avanti Polar Lipids.

Analysis of solution CD data. Measured values of ellipticity (θ) were converted into the ellipticity per amino acid residue [θ] by using

$$[\Theta] = \frac{\Theta M W_r}{lc}$$
[S3]

where l is the optical path-length of the cell, c the peptide concentration, and MW_R the average mass per amino acid residue of the peptide used. UV absorbance was measured with a Cary 3E spectrophotometer (Varian Analytical Instruments).

Molar concentrations were determined by using a molar extinction coefficient of $\varepsilon_{280 \text{ nm}} = 5600 \text{ M}^{-1} \text{ cm}^{-1}$ for tryptophan (a Trp was added to the C terminal of all peptides). Peptide concentrations were typically between 10 and 30 μ M in phosphate buffer.

Oriented CD measurements. The peptide and the lipid were codissolved in methanol at different molar ratios ($R_i = 1:12, 1:25$, 1:50, and 1:100) and carefully added drop-wise onto a 1-cmdiameter circular area in the center of a 2.5-cm-diameter quartz plate. After solvent removal under a stream of nitrogen followed by final removal under vacuum, samples were hydrated with warm air at 100% relative humidity. The quartz plate was mounted (sample side inward) on the end of a tube sealed at the opposite end with a second quartz plate. Samples were hydrated to their final hydration through the vapor phase by placing a drop of saturated salt solution in the sample holder before sealing. The sample was equilibrated overnight. The sample was then placed in a Jasco J720 CD spectrometer such that the optical axis was normal to the 2 parallel quartz plates. Spectra were recorded at each of 8 rotations of 45° around the optical axis, coinciding with the beam and normal to the lipid multilayers, and then averaged. The background signal was determined with the same amount of lipid, without the peptide.

All-Atom Molecular Dynamics Simulations. Using the Maestro molecular modeling environment (Schrödinger LLC), we built 2 sets of leucine α -helices (L_n) of varying length *n*, capped on each end with either weakly polar (GGPG) or positively charged (KKPK) groups in extended conformations. We first constructed 8 GGPG-L_n-GPGG peptides, the so-called GLn set, with n = 6, 8,10, 12, 14, 16, 18, and 20. We then mutated each glycine into lysine residues to generate 8 KKPK-L_n-KPKK peptides, the KLn set. Finally, we used psfgen (6) to acetylate the N termini and amidate the C termini of these 16 different systems. After applying strong harmonic constraints (100 kcal mol⁻¹) on the position of all heavy atoms, we used Nanoscale Molecular Dynamics (NAMD) software (6) to relax each peptide through a 1-ps NVT (constant number of particles, volume, and temperature) run in vacuo at 300 K.

We then aligned the principal moment of inertia of GL20 with the transmembrane axis of a POPC bilayer by positioning their centers-of-masses at the origin. Four lipids were removed to make room for the peptide. We applied strong harmonic constraints (100 kcal mol⁻¹) on the position of G20's backbone, and "froze" the rest of the system, except for an 8-Å-thick ring of

- Mouritsen OG, Bloom M (1993) Models of lipid-protein interactions in membranes. *Annu Rev Biophys Biomol Struc* 22:145–171.
- Kozak M (1999) Initiation of translation in prokaryotes and eukaryotes. Gene 234:187– 208.
- Sääf A, Wallin E, von Heijne G (1998) Stop-transfer function of pseudo-random amino acid segments during translocation across prokaryotic and eukaryotic membranes. *Eur J Biochem* 251:821–829.
- 5. Stewart JM, Young JD (1984) Solid Phase Peptide Synthesis (Pierce, Rockford, IL).
- 6. Kalé L, et al. (1999) NAMD2: Greater scalability for parallel molecular dynamics.
- J Comput Phys 151:283–312. 7. Tuckerman M, Berne BJ (1992) Reversible multiple time scale molecular dynamics. J Chem Phys 97(3):1990–2001.
- Grubmüller H, Heller H, Windemuth A, Schulten K (1991) Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Mol Simul* 6:121–142.
- Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. J Comput Phys 23:327–341.

lipids surrounding the peptide. We then minimized the system during 10,000 steps at 0 K, in absence of electrostatic interactions. We followed this with a 10-ps NVT run at 310 K, and then unfroze the lipids and the water molecules. We progressively removed constraints on the backbone from 100 to 20 to 5 to 0 kcal mol⁻¹ over three 1-ns NPT (constant number of particles, pressure, and temperature) runs (with electrostatic interactions) at 300 K and 1 bar. We then replicated the G20 system and proceeded to build the G18 system by replacing G20 with G18. We treated this new system in exactly the same way we relaxed the G20 system. After the peptide was unconstrained, we duplicated the G18 system to generate a starting point in which to insert G16, etc. Hence, we recursively built all systems down to G6. The KL20 system was built from the G20 system, to which we added 6 Cl⁻ counterions. Other KLn systems were generated recursively from K20, K18, etc.

Following peptide insertion into the bilayer, all MD simulations were performed with periodic boundary conditions by using a multiple time step integrator (7, 8) with an elementary time-step of 1 fs. Nonbonded and electrostatic interactions were calculated every 2 and 4 time-steps, respectively, with a cutoff of 11 Å. The SHAKE algorithm (9) was used to constrain the length of bonds involving hydrogen atoms. The particle mesh Ewald summation (10) was used in the calculation of Coulomb interactions. Water was modeled by using the TIP3P representation (11). The temperature was kept constant by using Langevin dynamics, and a Nosé-Hoover Langevin piston (12, 13) was used for pressure control (NPT, 1 bar and 300K). The MD simulations were performed with the NAMD program (6) using the all-atom CHARMM22 (14, 15) protein and CHARMM27 (16) lipid force fields. Molecular graphics images were prepared by using the Visual Molecular Dynamics (VMD) program (17).

We removed excess water from our systems to speed up simulations by reducing the aqueous content to 30 per lipid, which corresponds to the minimum water content of fully hydrated phosphatidylcholine bilayers (18). Hence, every system contained $n_{\text{lip}} = 280$ POPC molecules and 8,400 water molecules. Moreover, because of sheer computing cost and time constraints, we only retained the n = 6, 8, 10, and 12 systems, to study the effect of short helices on interactions, and the n = 20systems, to use as controls. We ran MD simulations on unconstrained systems until they were equilibrated, i.e., until their area per lipid was stable for 10 consecutive nanoseconds. The number of atoms, the size, the length of the equilibration run, and the area per lipid for each system are reported in Table 1.

- Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N●log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092.
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935.
- Martyna GJ, Tobias DJ, Klein ML (1994) Constant-pressure molecular-dynamics algorithms. J Chem Phys 101:4177–4189.
- 13. Feller SE, Zhang Y, Pastor RW, Brooks BR (1995) Constant pressure molecular dynamics simulation: The Langevin piston method. J Chem Phys 103:4613–4621.
- MacKerell AD, Jr, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102:3586–3616.
- MacKerell AD, Jr (2006) CHARMM Empirical Force Fields. Available at http:// www.pharmacy.umaryland.edu/faculty/amackere/force fields.htm. Accessed •••.
- Feller SE, MacKerell AD, Jr (2000) An improved empirical potential energy function for molecular simulations of phospholipids. J Phys Chem B 104:7510–7515.
- Humphrey W, Dalke W, Schulten K (1996) VMD: Visual molecular dynamics. J Mol Graphics 14:33–38.
- Nagle JF, Tristram-Nagle S (2000) Lipid bilayer structure. Curr Opin Struct Biol 10:474– 480.

Jensen MØ, Mouritsen OG (2004) Lipids do influence protein function: The hydrophobic matching hypothesis revisited. *Biochim Biophys Acta* 1666:205–226.



Fig. S1. CD spectra of synthetic peptides in methanol. Peptide concentrations are generally 40 μ M.



Fig. S2. CD spectra of synthetic peptides in 0.1 M phosphate buffer. Peptide concentrations were typically 25 μ M. (A) Spectra of GGPG-L_n-GPGG peptides. (B) Spectra of KKPK-L_n-KPKK peptides.



Fig. S3. Oriented CD data for the GGPG-Leu_n-GPGG (GL) and KKPK-Leu_n-KPKK (KL) peptides under various experimental conditions. The expected spectra for helices oriented normal and parallel to bilayers oriented normal to the CD spectrometer optical axis are shown in the *Upper Right* (GL6 in POPC).

S A Z d



Fig. S3. Continued.

S A Z d

SAZ

6 of 12



Fig. S3. Continued.

VAS

7 of 12



Fig. S3. Continued.

DNAS SANG

8 of 12



Fig. S4. Overview of molecular dynamics simulations. GGPG-L_n-GPGG (GLn) and KKPK-L_n-KPKK (KLn) systems for n = 6, 8, 10, 12, and 20. Every system contained 280 POPC molecules and 30 water molecules per lipid. Systems were ran at 300 K and 1 bar until their area per lipid was found to be stable for 10 consecutive nanoseconds. Information about each system is reported in Table 1. Color code: Green, glycine; blue, lysine; red, proline; and white, leucine. (*Upper*) G20 and K20 peptides in their respective POPC membrane before equilibration. Systems with shorter peptides were constructed from these bilayers. (*Lower*) All Gn and Kn peptides, averaged over the equilibrated runs. There are no major differences between the structure of the backbone of Gn and Kn systems. They all retained their helicity, except for K6.

DNA NG



Fig. S5. Some features of the GL6, GL20, and KL6 systems observed from a slice through the mean systems. (*Upper* and *Left*) A cut through the average GL20, GL6, and KL6 systems in van der Waals representation. The orange and light-blue lipid sections represent the polar headgroup and the apolar core, respectively. The acyl chains of GL6 and KL6 are strongly curled in an effort to accommodate the apolar helices. This is not the case of GL20 and KL20 (not shown here), around which the acyl chains are significantly straighter and more stretched. (*Lower* and *Right*) A cut through the KL6 backbone, surrounding water molecules, and the lipid in licorice representation, during the simulation. The KL6 peptide killed its helicity to maximize interactions between its capping groups and the bilayer's headgroup. Hence, intramolecular hydrogen bonds were replaced for interactions with water molecules, resulting in a trail of water going through the bilayer.



Fig. S6. Orientational order parameters for the GLn (a and b) and KLn (c and d) systems. Lipids are separated in 2 regions: Within 15 Å of the center of mass of the peptide's helix (region 1), lipids are less ordered than farther than 15 Å (region 2).



Fig. S7. Time-averaged number of hydrogen bonds in GL*n* and KL*n* systems as a function of the residue. Hydrogen-bonding occurs when 2 electro-negative atoms, usually nitrogen or oxygen, compete for the same hydrogen atom. The hydrogen atom is bonded covalently to the donor, D, and is being shared with the acceptor, A. In ideal cases, D-H-A form a 180° angle, but values can deviate by as much as 30° depending on the systems. Moreover, D and A are typically within 3 Å of each other. We used a purely geometrical procedure that is blind to the identity of the atoms interacting with each other, as long as D and A are within 3 Å, and as long as the D-H-A angle is within 150–210°. Nitrogen and oxygen are the only atoms that can participate in hydrogen-bonding in our systems. Our analysis has revealed that there was no occurrence of interactions in which nitrogen was an acceptor. Therefore, the only possibilities involve the backbone nitrogen as a donor and, as an acceptor : (*i*) backbone oxygen, (*ii*) phosphate oxygen, (*iii*) carbonyl oxygen, and (*iv*) water oxygen. In this figure, we labeled those combinations "backbone N as a donor, acceptor is oxygen, phosphate, carbonyl, and water". There are 4 more, nearly identical combinations in which the donor is the nitrogen from a lysine side chain instead. We labeled these "side chain N as a donor, acceptor is oxygens, and water". Moreover, we wanted to consider backbone oxygens as acceptors, and their interactions with peptide nitrogens and water acceptors. We labeled those combinations "side chain N as a donor, acceptor is nitrogen and water". The analysis is shown for (*a*) GL6, (*b*) GL20, (*c*) KL6, and (*d*) KL20. Information regarding the average number of hydrogen-bonding for all systems is presented in Table 1.

DNAS