Supporting Material for 'Thermodynamics of antimicrobial lipopeptide binding to membranes: Origins of affinity and selectivity'

Dejun Lin and Alan Grossfield*
Department of Biochemistry and Biophysics,
University of Rochester Medical Center,
Rochester, New York, U.S.A.

*Corresponding author. Address: Department of Biochemistry and Biophysics, University of Rochester Medical Center, 601 Elmwood Ave, Box 712, Rochester, New York 14620, U.S.A., Email: alan_grossfield@urmc.rochester.edu, Tel.: (585)276-4193

S1 Table of ΔG's

Table. S1 is a table of ΔG values as shown in Fig. 2.

	POPC	POPE:POPG
C16-KGGK	−11.6	−14.9
C16-KGGK (1M NaCl)	−11.8	−13.9
C16-GGGG	−11.2	−13.2
C16	−15.2	−15.0
KGGK	1.4	−0.7
GGGG	1.9	0.2
KGGK *	3.6	0.1
KGGK (1M NaCl) *	3.4	1.1
GGGG *	4.0	1.8

Table S1: Free energy of binding different ligands to POPC or POPE:POPG membrane in kcal/mol as shown in Fig. 2.

S2 Table of ΔΔG's

Table. S2 is a table of ΔΔG values as shown in Fig. 3.

	ΔΔG
C16-KGGK	−3.4
C16-KGGK (1M NaCl)	−2.1
C16-GGGG	−2.0
C16	0.2
KGGK	−2.1
GGGG	−1.7

Table S2: Difference in ΔG in kcal/mol of binding different ligands between the POPE:POPG membrane and POPC membrane as shown in Fig. 3.

*Value derived from ΔG(peptide) = ΔG(lipopeptide) - ΔG(C16)

S3 Boundary between the bound and unbound state

The boundary between the bound and unbound state is calculated from the distribution of the distance of the terminal hydrocarbon bead of C16-KGG**K** or C16-GGGG to membrane center at each umbrella window. These distributions are typically bimodal as shown in Fig. 4. The representative configurations of the two modes in this distribution are shown in Fig. S1. We found that there are 2 neighboring windows in each case between which the probability transitions rapidly to zero. The mean values of the centers of these 2 neighboring windows were chosen to be the boundary between the bound and unbound state. The list of boundaries in different system is listed in Table S3.

System	Boundary (Å)
C16-KGGK POPE:POPG	40.5
C16-KGGK POPC	38.5
C16-KGGK POPE:POPG (1M NaCl)	40.5
C16-KGGK POPC (1M NaCl)	38.5
C16-GGGG POPE:POPG	37.5
C16-GGGG POPC	35.5
C16 POPE:POPG	31
C16 POPC	29
KGGK POPE:POPG	30
KGGK POPC	28
GGGG POPE:POPG	27
GGGG POPC	25

Table S3: Boundary between the bound and unbound states of different systems.

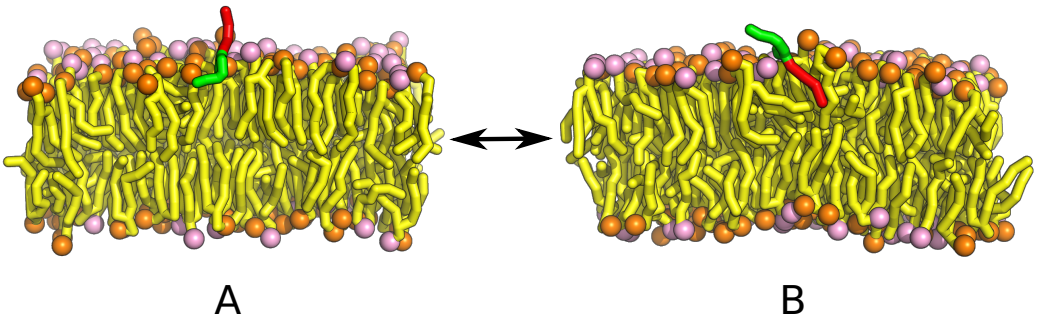

Figure S1: The representative configuration of C16-KGG**K** in a unbound (A) or bound (B) state on the surface of a POPE:POPG membrane. In either case, the lipids are rendered in yellow with the head groups of PE highlighted in orange and PG in pink. C16 is rendered in red while KGGK is in green.

S4 Comparison to equilibrium all-atom simulations

S4.1 All-atom (AA) simulations

The AA simulations were performed using the same protocol as described in our previous work (1). They were done using NAMD version 2.6 (2) with CHARMM22 force field (3) for protein and CHARMM27 force field for lipids (4–6) in the NPγT ensemble with constant particle number, pressure, surface tension and temperature. The surface tension, the pressure and the temperature was set at 27.5 dyn/cm, 1 atm and 300 K, respectively. We refer the readers to our previous paper (1) for more details.

The AA systems consist of 180 lipids symmetrically distributed across the two leaflets. As consistent with the coarse-grained (CG) simulations, two types of membranes were used, a pure POPC or a POPE:POPG with a ratio of 2:1. There are a total of 20 C16-KGGKs in each system with 10 of them on each leaflet. All the C16-KGGKs were fully inserted in the membrane. For type of membranes, we performed four independent simulations. All the analysis was based on the quantities averaged over the 4 simulations. We again refer the readers to our previous work (1) for details about the system construction. The analysis in this section was done using the software package LOOS (7).

S4.2 Center-of-mass distance (COM) distance between C16-KGGK and membrane center

For all-atom (AA) simulations, we computed the COM distance between each C16-KGG**K** and the membranes along the membrane normal from each frame of the simulations. The averages over all the frames of all the 4 simulations were 16.8 Å and 15.2 Å for POPE:POPG and POPC membranes, respectively.

S4.3 Lateral radial distribution functions (LRDFs)

The LRDFs of each lipids around C16-KGG**K** are defined as:

$$g(r_{xy,i}) = \frac{1}{N_{\text{pair}}\pi\left\{(r_{xy,i}+\frac{\delta}{2})^2-(r_{xy,i}-\frac{\delta}{2})^2\right\}}\left\langle n_{\text{pair}}(r_{xy,i})\right\rangle \tag{1}$$

where $r_{xy,i}$ is the distance in the plane of the membrane at the center of the ith bin, N_{pair} is the number of pairs possible (equal to N_aN_b if a and b are different chemical species, $N(N-1)/2$ if the RDF is for a single chemical species to itself), δ is the width of the histogram bins, and $n_{\text{pair}}(r_{xy,i})$ is the number of pairs found in distances belonging bin i in any given trajectory snapshot. We computed LRDFs from each simulation of the two systems and the averages over the four independent simulations for each system are reported in Fig. S2.

S4.4 Fractional contacts

We define the contacts that C16-KGG**K** make with other molecules, e.g., lipids, as the total number of heavy atoms (for AA simulations) or beads (for CG simulations) appear within a 5 Å radius centered on each particles of C16-KGG**K**. Such contacts are then divided by the total contacts that C16-KGG**K** makes with other C16-KGG**K**, lipids and water to give the fractional contacts. Again, the averages over the four simulations are shown in Fig. S3.

S5 Projecting the potentials of mean force on the peptide-membrane degree of freedom

Consider two continuous variable x and y, representing two degrees of freedom in the system whose potentials of mean force (PMFs) we are interested in. The PMF as a function of (x, y), $\omega(x, y)$, is related to the respective probability distribution function by

$$\omega(x,y) \equiv -k_BT\ln\rho(x,y) \tag{2}$$

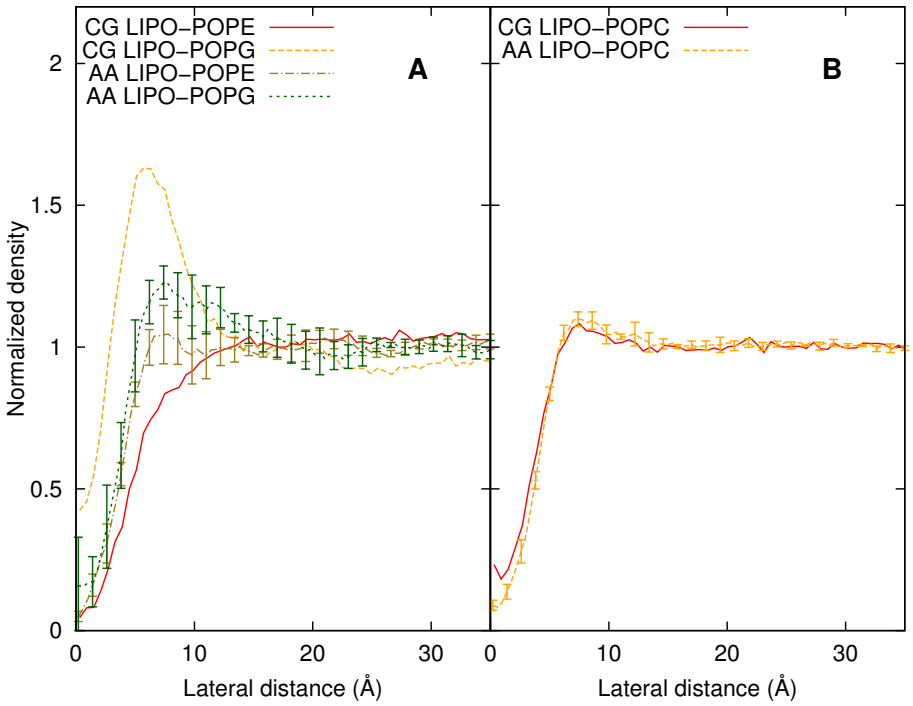

Figure S2: The lateral radial distribution functions (LRDFs) of different lipids around C16-KGG**K** in both coarse-grained (CG) and all-atom (AA) simulations. The CG simulations are the umbrella sampling windows corresponding the PMF minima. Panel *A* and *B* are computed from simulations using POPE:POPG and POPC membranes respectively. The AA data were the averages over 4 simulations and the error bars are the standard deviation.

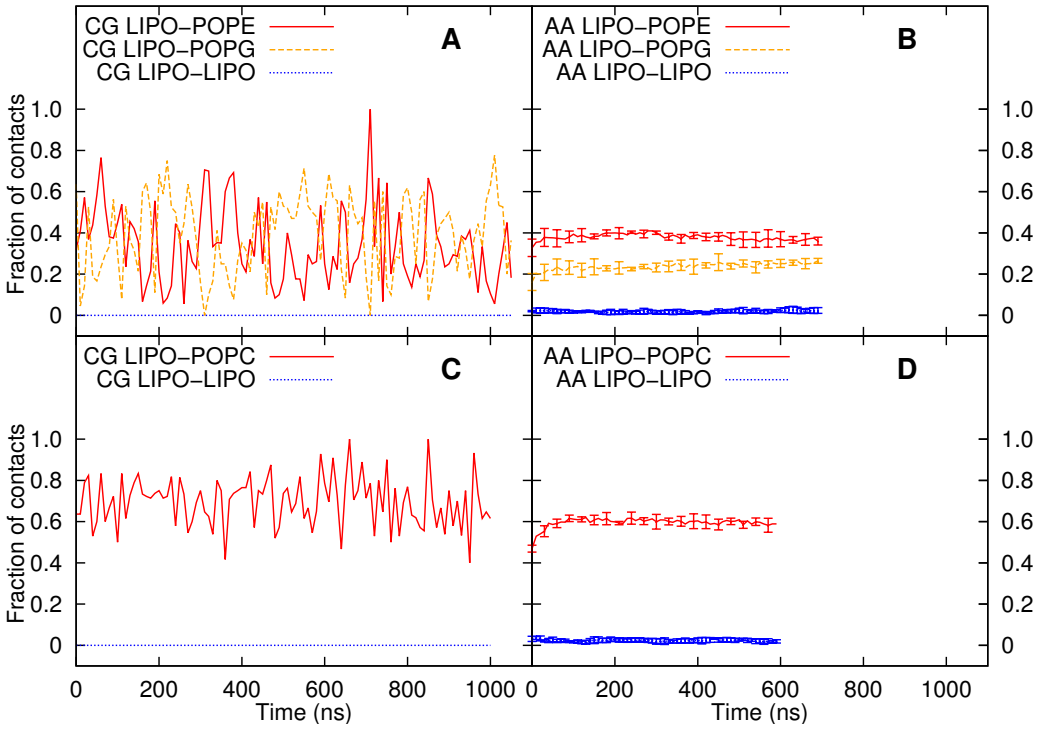

Figure S3: The fractional contacts C16-KGG**K** made with other C16-KGG**K**, different lipids and water as a function of simulation time. Panel *A* and *C* are from coarse-grained (CG) umbrella sampling windows corresponding the PMF minima while *B* and *D* are from all-atom (AA) simulations. Panel *A* and *B* are computed from simulations using POPE:POPG membranes while *C* and *D* are from those using POPC membranes. For clarity, we omitted the plot of water contacts. The AA data were the averages over 4 simulations and the error bars are the standard deviation.

and

$$\rho(x, y) = \rho(y|x)\rho(x) \tag{3}$$

where $\rho(y|x)$ is the conditional probaiblity density function of y given x. We carried out a set of umbrella sampling simulations with restraints applied on x only and obtained an estimate of $\rho(x)$ from the Weighted Histogram Analysis Method (WHAM): $\hat{\rho}_x$, where we denote the value of the histogram-based estimator at the grid point centered at x by the corresponding subscripted symbol. The estimated values are denoted with a circumflex to distinguish them from the underlying true ones. From the same set of simulations, we also obtained the estimate of $\rho(y|x)$ by histogramming (x, y) on a grid:

$$\hat{\rho}(y|x) = \hat{\rho}'_{x,y}/\hat{\rho}'_x \tag{4}$$

where $\hat{\rho}'_{x,y}$ is just the total histogram value (if we assume uniform grid size on each variable) at grid point centered at (x, y), summed over all the histograms at the same grid point from all the restrained simulations. Such histograms are denoted primed since they are obtained from biased simulations to distinguish them from the corresponding unbiased ones. Similarly, $\hat{\rho}'_x \equiv \sum_y \hat{\rho}'_{x,y}$. Thus, we have the unbiased estimate of $\rho(x, y)$ as

$$\hat{\rho}_{x,y} = \hat{\rho}_x\hat{\rho}'_{x,y}/\hat{\rho}'_x \tag{5}$$

and PMF on y as

$$\hat{\omega}(y) = -k_BT \ln \sum_x \hat{\rho}_{x,y} \tag{6}$$

The estimates in equations 5 and 6 are optimal and can be derived from the original WHAM equation (8).

Applying equations 5 and 6 to our system, where x is the center-of-mass (COM) distance between the whole lipopeptide and the membrane along the membrane normal and y is the corresponding distance between the peptide and the membrane. These PMFs $\hat{\omega}(y)$ are shown in Fig. S4 together with $\hat{\omega}(x)$ for comparison.

S6 Potential of mean forces calculation using the polariable MARTINI water model

The systems were identical to the unpolarizable counterparts except that the

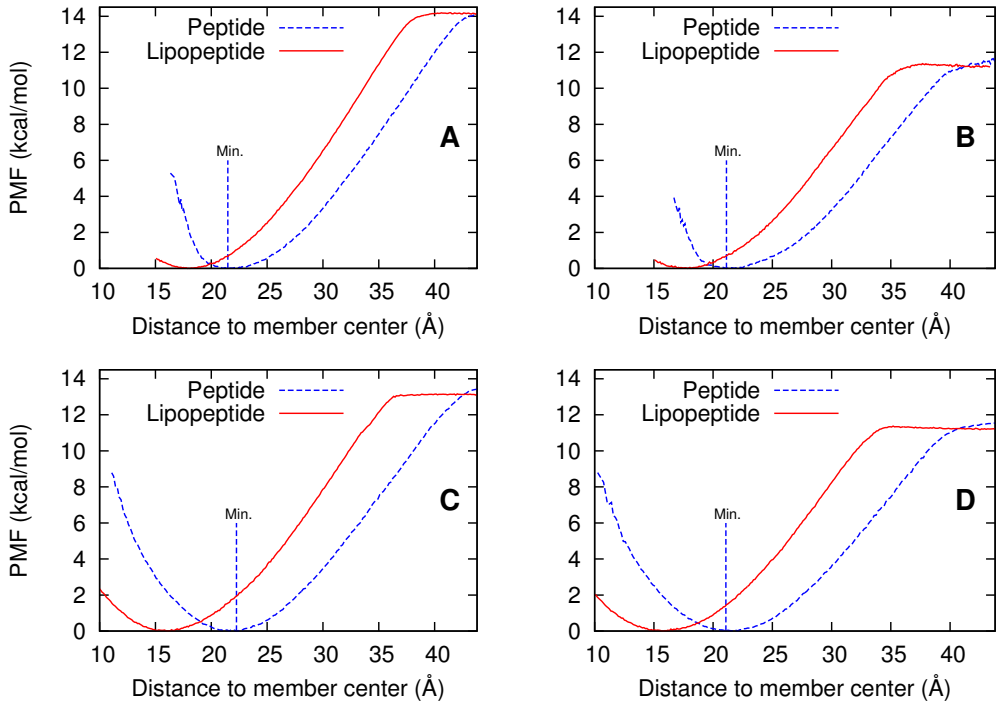

Figure S4: Potentials of mean force (PMFs) in kcal/mol as a function of the center-of-mass (COM) distance between either the lipopeptides (solid lines) or the peptides (dashed lines) and the membranes. The minimum of each peptide PMF is denoted by a vertical line on each graph. Panel *A*: C16-KGGK with POPE:POPG, *B*: C16-KGGK with POPC, *C*: C16-GGGG with POPE:POPG, *D*: C16-GGGG with POPC.

unpolarizable water particles were replaced by the same number of the polarizable ones using the script provided on the MARTINI website (`http://md.chem.rug.nl/cgmartini/images/tools/water2polarizable/triple-w.py`). This dramatically increased the number of particles in the simulations, from 21385 to 50899. Each system after the water replacement was subjected to energy minimization followed by at least 1 ns of equilibrating MD simulations before the production umbrella sampling run (see below). The force field version 2.2P (last modified on 08-8-2012) (9, 10) was used in the simulations.

Umbrella sampling with Hamiltonian replica exchange (HREX) was used

to calculate the potential of mean forces (PMFs) as a function of the same reaction coordinate (RC) as we used in the unpolarizable water model (see section 2). The HREX approach has been shown to facilitate the convergence of the PMF (11, 12), and we can afford better convergence with less simulation time as compared to standard umbrella sampling. The range of the reaction coordinate from 15 Å to 50 Å was divided into 48 evenly spaced windows. Harmonic restraints were applied to the RC in each window with a force constant of 1000 kJ/(mol nm^{-2}).

A total of 48 simulations, with one simulation for each umbrella window were performed using GROMACS 4.6.3 (13, 14) with modification so that we can have the center of the harmonic restraints exchanged between different windows. The simulation control parameters were the same as in the unpolarizable water case 2 except that the relative dielectric constant is reduced from 15 to 2.5 (9). The HREX was attempted every 500 steps. The details of the HREX algorithm were described in this reference (11). Each replica/simulation was run for 400 ns where the first 100 ns of simulation data were treated as equilibration phase and excluded from further analysis. The calculation of the PMFs and the binding free energy followed the same protocol as described in section 2.

Supporting References

1. Horn, J. N., T. D. Romo, and A. Grossfield, 2013. Simulating the mechanism of antimicrobial lipopeptides with all-atom molecular dynamics. *Biochemistry* 52:5604–5610. http://dx.doi.org/10.1021/bi400773q.

2. Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kal, and K. Schulten, 2005. Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802. http://dx.doi.org/10.1002/jcc.20289.

3. MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wirkiewicz-Kuczera, D. Yin, and M. Karplus, 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616. http://dx.doi.org/10.1021/jp973084f.

4. Feller, S. E., and A. D. MacKerell, 2000. An Improved Empirical Potential Energy Function for Molecular Simulations of Phospholipids. *J Phys Chem B* 104:7510–7515. http://pubs.acs.org/doi/abs/10.1021/jp0007843.

5. Klauda, J. B., B. R. Brooks, A. D. MacKerell, Jr, R. M. Venable, and R. W. Pastor, 2005. An ab initio study on the torsional surface of alkanes and its effect on molecular simulations of alkanes and a DPPC bilayer. *J Phys Chem B* 109:5300–5311. http://dx.doi.org/10.1021/jp0468096.

6. Feller, S. E., K. Gawrisch, and A. D. MacKerell, Jr, 2002. Polyunsaturated fatty acids in lipid bilayers: intrinsic and environmental contributions to their unique physical properties. *J Am Chem Soc* 124:318–326.

7. Romo, T. D., and A. Grossfield, 2009. LOOS: an extensible platform for the structural analysis of simulations. *Conf Proc IEEE Eng Med Biol Soc* 2009:2332–2335. http://dx.doi.org/10.1109/IEMBS.2009.5335065.

8. Kumar, S., J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, 1992. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13:1011–1021. http://dx.doi.org/10.1002/jcc.540130812.

9. Yesylevskyy, S. O., L. V. Schäfer, D. Sengupta, and S. J. Marrink, 2010. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput Biol* 6:e1000810. http://dx.doi.org/10.1371/journal.pcbi.1000810.

10. de Jong, D. H., G. Singh, W. F. D. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schfer, X. Periole, D. P. Tieleman, and S. J. Marrink, 2013. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J Chem Theory Comput* 9:687–697. http://pubs.acs.org/doi/abs/10.1021/ct300646g.

11. Sugita, Y., A. Kitao, and Y. Okamoto, 2000. Multidimensional replica-exchange method for free-energy calculations. *J Chem Phys* 113:6042–6051. http://scitation.aip.org/content/aip/journal/jcp/113/15/10.1063/1.1308516.

12. Woods, C. J., J. W. Essex, and M. A. King, 2003. The Development of Replica-Exchange-Based Free-Energy Methods. *J Phys Chem B* 107:13703–13710. http://pubs.acs.org/doi/abs/10.1021/jp0356620.

13. Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl, 2008. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* 4:435–447. http://pubs.acs.org/doi/abs/10.1021/ct700301q.

14. Van Der Spoel, D., E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, 2005. GROMACS: fast, flexible, and free. *J Comput Chem* 26:1701–1718. http://dx.doi.org/10.1002/jcc.20291.