Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Electronic Supplementary Information: Thermodynamics of Cell Penetrating Peptides on Lipid Membranes: Sequence and Membrane Acidity Regulate Surface Binding

Pedro G. Ramírez,[†] Mario G. Del Pópolo,[‡] Jorge A. Vila,[†] and Gabriel S. Longo[*,¶]

†*Instituto de Matemática Aplicada San Luis (IMASL), UNSL-CONICET, San Luis, Argentina.*

‡*Instituto Interdisciplinario de Ciencias Básicas (ICB-CONICET) & Facultad de Ciencias Exactas y Naturales (FCEN-UNCuyo), Mendoza, Argentina.*

¶*Instituto de Investigaciones Fisicoquímicas, Teóricas y Aplicadas (INIFTA), UNLP-CONICET, La Plata, Argentina.*

E-mail: longogs@inifta.unlp.edu.ar

Contents

1 Methods

1.1 Molecular Theory

Let us consider a lipid membrane separating two regions of an aqueous solution that contains peptide molecules (CPP), water molecules (w), hydroxyde ions (OH^-), hydronium ions (OH_3^+ or H^+), and monovalent salt anions ($-$) and cations ($+$). Sufficiently far from the membrane surface, the chemical composition of this solution is experimentally controlled and completely defined by its pH, salt concentration, [NaCl], and peptide concentration, [CPP]. The membrane is composed of two different types of lipids; a charge neutral species (N) and another bearing an acidic group in its head-group (I). This ionizable head-group can be found in one of two possible chemical states: either protonated (charge neutral, IH) or deprotonated (negatively charged, I^-).

We want to study the equilibrium properties of this system, including peptide adsorption to the membrane surface and lipid charge regulation. Therefore, we express the Helmholtz free energy of the system as

$$F = -TS_{TF} - TS_{TP} + F_{CP} + F_{MS} + U_E + U_{AD} + U_{st} \tag{S1}$$

where

- T is the temperature.
- S_{TF} is the translational entropy of free species (excluding the peptide): water and small ions.
- S_{TP} is the translational and conformational entropy of peptides.
- F_{CP} is the chemical free energy of peptides (acid-base equilibrium).

- F_{MS} is the surface (non-electrostatic) free energy of the lipid membrane, which includes description of the acid-base equilibrium of ionizable lipids.
- U_E is the electrostatic energy.
- U_{AD} is the contribution that accounts for the internal energy of contacts between the membrane surface and peptide residues.
- U_{st} is the internal energy of the steric interactions.

The membrane interior is modeled as a region from which solution molecules are excluded, having permittivity ϵ_M, different from that of the solvent (ϵ_w). This region is defined by

$$-h_{mem} < z < 0 \tag{S2}$$

where h_{mem} is the membrane thickness, and z is the distance from the planar upper membrane surface, defined by the plane $z = 0$. The lower membrane surface, given by $z = -h_{mem}$, is also planar. We assume there is reflection symmetry with respect to the $z = -\frac{h_{mem}}{2}$ plane. Next, we will present expressions for the different contributions to the free energy Eq. (S1).

The translational entropy of the small free species is given by

$$-\frac{S_{TF}}{k_B} = \sum_{\gamma \in \{w, OH_3^+, OH^-, +, -\}} A \int_0^\infty dz\, \rho_\gamma(z) \left[\ln\left(\rho_\gamma(z) v_w\right) - 1 + \beta \mu_\gamma^0 \right] \tag{S3}$$

where

- γ runs over the different small species in the solution.
- $\rho_\gamma(z)$ is the number density of species γ at z; namely, the number of these molecules between z and $z + dz$.
- μ_γ^0 is the standard chemical potential of species γ. Then, in addition to the mixing/translation entropy, S_{TF} also incorporates the self-energies of solution species.

- v_w is the volume of a water molecule, included to keep the inside of the logarithms unitless.
- k_B is the Boltzmann constant and $\beta = \frac{1}{k_B T}$.
- A is the area of the membrane surface.

The entropy of peptide translations and conformations can be expressed as

$$-\frac{S_{TP}}{k_B} = A \int_0^\infty dz \sum_{\alpha_P \in \{\alpha_P\}} \rho_P(\alpha_P, z) \left[\ln \left(\rho_P(\alpha_P, z) v_w \right) - 1 + \beta \mu_P^0 \right] \tag{S4}$$

where

- α_P denotes a peptide conformation; $\{\alpha_P\}$ is the set of all peptides conformations. If the peptide does not have conformational flexibility, then $\{\alpha_P\}$ simply contains rotations of the rigid structure. On the contrary, $\{\alpha_P\}$ contains both different conformers and rotations.
- $\rho_P(\alpha_P, z)$ is the local number density of peptides in conformation α_P. The total density of peptide is an ensemble average over its configurations:

$$\langle \rho_P(z) \rangle = \sum_{\alpha_P} \rho_P(\alpha_P, z) \tag{S5}$$

- μ_P^0 is the standard chemical potential of peptides, which also considers the self-energy of the molecule. If peptide conformations have different intrinsic energies associated, then $\mu_P^0 = \mu_P^0(\alpha_P)$.
- We can now define the probability distribution function of molecular conformations, which depends on the position of the center of mass of the peptide:

$$P_P(\alpha_P, z) = \frac{\rho_P(\alpha_P, z)}{\langle \rho_P(z) \rangle} \tag{S6}$$

which satisfies

$$\sum_{\alpha_P} P_P(\alpha_P, z) = 1 \; ; \forall \; z \tag{S7}$$

In this work, we assume that all peptide conformations are equally probable in the bulk solution. Namely,

$$\lim_{z\to\infty} P_P(\alpha_P, z) = \frac{1}{n_c} \tag{S8}$$

where n_c is the total number of peptide conformations; $n_c = \dim\{\alpha_P\}$.

The chemical free energy of peptides, which describes the acid-base equilibrium of titratable residues can be written as

$$\begin{aligned}\beta F_{CP} = A\int_0^\infty dz \sum_{\tau\in\{\tau\}} \langle\rho_\tau(z)\rangle g_\tau(z)\big[\ln\big(g_\tau(z)\big) + \beta\mu^0_{\tau,p}\big] \\ + A\int_0^\infty dz \sum_{\tau\in\{\tau\}} \langle\rho_\tau(z)\rangle (1-g_\tau(z))\big[\ln\big(1-g_\tau(z)\big) + \beta\mu^0_{\tau,d}\big]\end{aligned} \tag{S9}$$

where

- τ denotes a particular type of protonable unit; $\{\tau\}$ is the whole set of different such units. In this work, we will use the terms unit/segment and residue interchangeably. However, this present formulation of the theory does not require amino-acid residues to be described with a single coarse-grain unit.
- $\langle\rho_\tau(z)\rangle$ is the local number density of τ segments, which is an ensemble average over peptide conformations:

$$\langle\rho_\tau(z)\rangle = A\int_0^\infty dz' \sum_{\alpha_P\in\{\alpha_P\}} \rho_P(\alpha_P, z') n_\tau(\alpha_P, z', z) \tag{S10}$$

- $n_\tau(\alpha_P, z', z)$ is the number of τ units that a single peptide in conformation α_P and center of mass at z' contributes to z.
- To calculate results using the theoretical framework described in the section, a

molecular model must be defined that provides $\{n_\tau(\alpha_P, z', z), \forall(\alpha_P, z', z)\}$ as an input for each peptide.

- Note that Eq. (S10) is valid for all units, not only those that are titratable, including charge neutral units.
- $g_\tau(z)$ is the local degree of protonation of a τ unit. Then,

$$g_\tau(z)\langle\rho_\tau(z)\rangle \tag{S11}$$

gives the number of protonated τ-units between z and $z + dz$; similarly,

$$(1 - g_\tau(z))\langle\rho_\tau(z)\rangle \tag{S12}$$

is the local density of these units in the deprotonated state.

- $\mu^0_{\tau,p}$ y $\mu^0_{\tau,d}$ are respectively the standard chemical potentials of the protonated and deprotonated states of type-τ segments.

The surface free energy accounts for the chemical energy of the charged and protonated acidic lipids and the mixture between them. This contribution can be expressed as

$$\begin{aligned}\frac{\beta F_{MS}}{A} &= \sigma_I f_{I^-}(\ln f_{I^-} + \beta\mu^0_{I^-}) \\ &\quad + \sigma_I(1 - f_{I^-})\left[\ln(1 - f_{I^-}) + \beta\mu^0_{IH}\right]\end{aligned} \tag{S13}$$

where

- σ_I is the surface density of ionizable lipid head-groups; similarly, σ_N is that of neutral lipids.
- $\mu^0_{I^-}$ and μ^0_{IH} are the standard chemical potentials of the dissociated (charged) and protonated ionizable lipid, respectively.

- f_{I^-} is the local degree of dissociation of the ionizable lipid, such that

$$\sigma_{I^-} = f_{I^-}\sigma_I \tag{S14}$$

is the area density of charged lipid, while

$$\sigma_{IH} = (1 - f_{I^-})\sigma_I \tag{S15}$$

is the area density of uncharged ionizable lipid.

- The membrane-solution interface is fully occupied by lipid head-groups, which means

$$\sigma_I a_I + \sigma_N a_N = 1 \tag{S16}$$

where a_I is the area occupied by a ionizable lipid (on the interface), and a_N is that of a single neutral lipid. Note that the fraction of area occupied by ionizable lipid is $x_I = \sigma_I a_I$.

The electrostatic energy can be written as

$$\begin{aligned}\frac{\beta U_E}{A} = &\int_{-h_{mem}/2}^{\infty} \left[\langle \rho_q(z) \rangle \beta\psi(z) - \frac{1}{2}\beta\epsilon(z) \left(\frac{\partial \psi(z)}{\partial z} \right)^2 \right] dz \\ &+ \sigma_I f_{I^-} q_{I^-} \psi_S\end{aligned} \tag{S17}$$

where

- $\langle \rho_q(z) \rangle$ is the local charge density:

$$\langle \rho_q(z) \rangle = \sum_\gamma \rho_\gamma(z) q_\gamma + \sum_\tau q_\tau f_\tau(z) \langle \rho_\tau(z) \rangle \tag{S18}$$

- $f_\tau(z)$ is the local charge degree of τ units, such that the number density of charged

τ units at position z is

$$f_\tau(z)\langle\rho_\tau(z)\rangle \tag{S19}$$

while the number density of uncharged τ units is

$$(1 - f_\tau(z))\,\langle\rho_\tau(z)\rangle \tag{S20}$$

Note that $f_\tau(z)$ is related to the degree of protonation $g_\tau(z)$:

$$f_\tau(z) = \begin{cases} 1 - g_\tau(z) & \text{for acidic } \tau \text{ units} \\ g_\tau(z) & \text{for basic } \tau \text{ units} \end{cases} \tag{S21}$$

- q_γ is the electric charge of the free species γ, q_τ is that of charged titratable units of the CPP, and q_{I^-} corresponds to the ionized lipid head -groups.
- $\psi\,(z)$ is the local electrostatic potential, where

$$\psi_S = \psi(z = 0) \tag{S22}$$

is its value at the membrane-solution interface.

- $\epsilon\,(z)$ if the medium permittivity, which is

$$\epsilon(z) = \begin{cases} \epsilon_w & \text{in the solution} \\ \epsilon_M & \text{inside the membrane} \end{cases} \tag{S23}$$

The internal energy of lipid-peptide contacts, U_{AD}, can be expressed as:

$$\beta U_{AD} = A \int_0^\infty dz \sum_{\alpha_P} \rho_p(\alpha_P, z) \sum_\eta \beta\chi_\eta n_\eta\left(\alpha_P, z, z' = 0\right) \tag{S24}$$

where

- χ_η is the energy of a single contact between a η unit and the membrane surface; η runs over all peptide units. In this work, we model each amino-acid residue as a single coarse-grain particle, and assign $\chi_\eta = 0$ for all residues except tryptophan ($\chi_W \neq 0$).

- $n_\eta\left(\alpha_P, z, z' = 0\right)$ is the number of η residues in contact with the membrane surface, for a peptide in conformation α_P and center of mass at z. More generally, $n_\eta\left(\alpha_P, z, z'\right)$ has been defined in the context of Eq. (S10).

Th steric interactions of U_{st} are included at the at the excluded volume level through the incompressibility constraint that the equilibrium conditions must satisfy. The solution is incompressible, which means that every element of volume must be completely occupied by a combination of the different molecular species; this is

$$\sum_{\gamma\in\{w,H^+,OH^-,+,-\}} \rho_\gamma(z) v_\gamma + \langle \phi_P(z)\rangle = 1 \tag{S25}$$

for all $z > 0$, where

- v_γ is the molecular volume of species γ.
- $\langle \phi_P(z)\rangle$ is the total volume fraction occupied by peptide units:

$$\langle \phi_P(z)\rangle = \sum_{\eta\in\{\eta\}} \langle \rho_\eta(z)\rangle v_\eta \tag{S26}$$

  where again η runs over all peptide units, including titratable and non-titratable ones. Namely, the set $\{\tau\}$ is a subset of $\{\eta\}$.

- v_η is the molecular volume of the peptide's η unit.

In addition, another constraint that the free energy (Eq. (S1)) must satisfy is

$$A\int_0^\infty \left[\sum_\gamma \rho_\gamma(z)\, q_\gamma + \sum_\tau q_\tau f_\tau(z)\,\langle\rho_\tau(z)\rangle\right] dz + A\sigma_I q_{I^-} f_{I^-} = 0 \tag{S27}$$

which imposes global electroneutrality of the system.

After presenting expressions for all its contributions, the free is a *functional* of a few functions, which remain to be determined:

i. $\rho_\gamma(z)$, the local number density of free species γ (for all γ's).

ii. $\rho_P(\alpha_P, z)$, the local number density of each peptide conformation.

iii. $g_\tau(z)$ (or $f_\tau(z)$), the local degree of protonation (charge) of peptide titratable units.

iv. f_{I^-}, the degree of charge of ionizable lipid head-groups.

v. $\psi(z)$, the local electrostatic potential, including ψ_S.

Any other quantity is either an input of the theory or it can be calculated using functions (i) to (v). To obtain expressions for these functions (i-v), we optimize the proper thermodynamic potential with respect to each of them.

For this system, the thermodynamic equilibrium corresponds to the minimum of the semi-grand potential, Ω, which is function of the chemical potentials of all the solution species. This is because the system is in chemical equilibrium with a bulk solution of controlled composition. Thus,

$$\Omega = F - \sum_{\gamma\in\{H^+,OH^-,+,-\}} \mu_\gamma N_\gamma - \mu_P N_P \tag{S28}$$

where

- μ_γ, μ_P are the chemical potentials of free species γ and the peptide, respectively;

N_γ y N_P, represent the corresponding number of molecules.

- The sum over γ in Eq. (S28) excludes water because the incompressibility constraint, Eq. (S25), reduces the number of independent variables. In addition, because of Eq. (S16) the number of ionizable lipid molecules, $N_I = \sigma_I A$, and that of neutral lipids, $N_N = \sigma_N A$, are not independent from each other. Namely,

$$\Omega(T, V, A, N_N, N_I, \mu_w, \mu_{H^+}, \mu_+, \mu_-, \mu_P) \equiv \Omega(T, V, A, N_I, \mu_{H^+}, \mu_+, \mu_-, \mu_P) \tag{S29}$$

  where the μ's on the right-hand side of this equation are exchange chemical potentials (likewise in Eq. (S28)).

Equation (S28) can be expressed as

$$\begin{aligned}\Omega = F &- \sum_{\gamma\in\{H^+,OH^-,+,-\}} \mu_\gamma \left(A \int_0^\infty dz\, \rho_\gamma(z) \right) \\ &- \mu_P \left(A \int_0^\infty dz \sum_{\alpha_P\in\{\alpha_P\}} \rho_P(\alpha_P, z) \right) \\ &- \mu_{H^+} \left[A \int_0^\infty dz \sum_{\tau\in\{\tau\}} g_\tau(z) \langle \rho_\tau(z) \rangle + A\sigma_I (1 - f_{I^-}) \right]\end{aligned} \tag{S30}$$

where the last term has been introduced to properly count protons, including free hydronium ions ($\rho_{H^+}(z)$), and those that protonate the titratable units of lipid headgroups and peptides. In other words, these terms provide the proper N_{H^+} that is conjugated to μ_{H^+}.

To explicitly include the restriction of solution incompressibility, Eq. (S25), in the function to optimize, we define:

$$\beta\Phi = \beta\Omega + A \int_0^\infty dz\, \beta\pi_v(z) \left(\sum_{\gamma\in\{w,H^+,OH^-,-,+\}} \rho_\gamma(z) v_\gamma + \langle \phi_P(z) \rangle - 1 \right) \tag{S31}$$

where

- $\pi_v(z)$ is the local Lagrange multiplier introduced to reinforce the constraint at each position.

Explicitly, the function to optimize is

$$
\begin{aligned}
\beta\Phi = & \sum_{\gamma\in\{w,H^+,OH^-,+,-\}} A\int_0^\infty dz\,\rho_\gamma(z)\left[\ln\left(\rho_\gamma(z)v_w\right)-1+\beta\mu_\gamma^0\right] \\
& + A\int_0^\infty dz \sum_{\alpha_P\in\{\alpha_P\}} \rho_P(\alpha_P,z)\left[\ln\left(\rho_P(\alpha_P,z)v_w\right)-1+\beta\mu_P^0\right] \\
& + A\int_0^\infty dz \sum_{\tau\in\{\tau\}} \langle\rho_\tau(z)\rangle g_\tau(z)\left[\ln\left(g_\tau(z)\right)+\beta\mu_{\tau,p}^0\right] \\
& + A\int_0^\infty dz \sum_{\tau\in\{\tau\}} \langle\rho_\tau(z)\rangle (1-g_\tau(z))\left[\ln\left(1-g_\tau(z)\right)+\beta\mu_{\tau,d}^0\right] \\
& + A\sigma_I f_{I^-}(\ln f_{I^-}+\beta\mu_{I^-}^0) \\
& + A\sigma_I(1-f_{I^-})\left[\ln\left(1-f_{I^-}\right)+\beta\mu_{IH}^0\right] \\
& + A\int_{-h_{mem}/2}^\infty dz\left[\langle\rho_q(z)\rangle\beta\psi(z)-\frac{1}{2}\beta\epsilon(z)\left(\frac{\partial\psi(z)}{\partial z}\right)^2\right] \\
& + A\sigma_I f_{I^-} q_I\beta\psi_S \\
& - \sum_{\gamma\in\{H^+,OH^-,+,-\}} \beta\mu_\gamma\left(A\int_0^\infty dz\,\rho_\gamma(z)\right) \\
& - \beta\mu_{H^+}\left[A\int_0^\infty dz \sum_{\tau\in\{\tau\}} g_\tau(z)\langle\rho_\tau(z)\rangle+\sigma_I(1-f_{I^-})\right] \\
& - \beta\mu_P\left(A\int_0^\infty dz \sum_{\alpha_P\in\{\alpha_P\}} \rho_P(\alpha_P,z)\right) \\
& + A\int_0^\infty dz\,\beta\pi_v(z)\left(\sum_{\gamma\in\{w,H^+,OH^-,-,+\}} \rho_\gamma(z)v_\gamma+\langle\phi_P(z)\rangle-1\right)
\end{aligned}
\tag{S32}
$$

Note that we do not define a Lagrange multiplier to reinforce global electroneutrality, Eq. (S27); this condition is achieved through (a) imposing the proper boundary conditions to the solution of the Poisson Equation, and (b) considering the electroneutrality of the bulk solution in chemical equilibrium with our system.

Optimization with respect to $\rho_\gamma(z)$ leads to

$$\rho_\gamma(z) = \frac{e^{\beta\mu_\gamma - \beta\mu_\gamma^0}}{v_w} \exp\big(-\beta\pi_v(z)v_\gamma - \beta\psi(z)q_\gamma\big) \tag{S33}$$

which is valid for all small free species, including water with $e^{\beta\mu_w - \beta\mu_w^0} = 1$ and $q_w = 0$. The quantity

$$a_\gamma = \exp\left(\beta\mu_\gamma - \beta\mu_\gamma^0\right) \tag{S34}$$

is the activity of the species. All activities are completely determined once the composition of the bulk solution is given (pH, salt concentration, and peptide concentration). Indeed, we can write the number density of species γ as

$$\rho_\gamma(z) = \frac{\rho_\gamma^b}{(v_w\rho_w^b)^{\frac{v_\gamma}{v_w}}} \exp\big(-\beta\pi_v(z)v_\gamma - \beta\psi(z)q_\gamma\big) \tag{S35}$$

where ρ_γ^b is the number density in the homogeneous bulk solution (ρ_w^b being that of water).

Optimization with respect to $f_\tau(z)$ leads to

$$\frac{f_\tau(z)}{1-f_\tau(z)} = \frac{f_\tau^b}{1-f_\tau^b} e^{-\beta\psi(z)q_\tau} \tag{S36}$$

where

- f_τ^b is the degree of charge of the corresponding unit in the *bulk*, which is completely determined by the intrinsic pKa of the isolated unit and the bulk solution pH:

$$\frac{f_\tau^b}{1-f_\tau^b} = \begin{cases} \frac{K_\tau^0}{a_{H^+}} & \text{for acidic } \tau \text{ units} \\ \frac{a_{H^+}}{K_\tau^0} & \text{for basic } \tau \text{ units} \end{cases} \tag{S37}$$

- K_τ^0 is the thermodynamic equilibrium constant of the acid-base reaction of τ units, which satisfies

$$K_\tau^0 = e^{\beta\mu_{\tau,p}^0 - \beta\mu_{\tau,d}^0 - \beta\mu_{H^+}^0} \tag{S38}$$

Optimization with respect to $\rho_P(\alpha_P, z)$ leads to

$$\rho_P(\alpha_P, z) = x_P \exp\left(- A \int_0^\infty dz' M_P(\alpha_P, z, z') \right) \tag{S39}$$

with

$$\begin{aligned} M_P(\alpha_P, z, z') = & \sum_{\tau\in\{\tau\}} n_\tau(\alpha_P, z, z') \Big(\ln f_\tau(z') + \beta\psi(z') \Big) \\ & + \sum_{\eta\in\{\eta\}} n_\eta(\alpha_P, z, z') \beta\pi_v(z') v_\eta \end{aligned} \tag{S40}$$

and

$$x_P = e^{\beta\mu_P - \beta\mu_P^0} \prod_{\tau\in\{\tau\}} \left(\frac{g_\tau^b}{1 - f_\tau^b} e^{-\beta\mu_{\tau,d}^0} \right)^{N_\tau} = \frac{\rho_P^b(\alpha_P)}{(\rho_w^b)^{\frac{v_P}{v_w}}} \prod_{\tau\in\{\tau\}} (f_\tau^b)^{N_\tau} \tag{S41}$$

where

- $\rho_P^b(\alpha_P)$ is the density of peptide conformer α_P in the bulk solution.
- g_τ^b y f_τ^b are degrees of protonation and charge in the bulk solution, respectively.
- N_η (or N_τ) is the total number of η units (or τ) in a peptide molecule:

$$N_\eta = \int_0^\infty dz'\, n_\eta(\alpha_P, z, z') \tag{S42}$$

  independently of (α_P, z), provided that the conformation fits in the system (*i.e.*, it does not protrude the membrane when the peptide places its center of mass at z); otherwise, $n_\eta(\alpha_P, z, z') \equiv 0\, \forall z'$ for this particular α_P (other conformations might still fit at z).

- v_P is the total volume of a peptide molecule:

$$v_P = \sum_{\eta \in \{\eta\}} N_\eta v_\eta \tag{S43}$$

  where v_η is the volume of a η unit.

- As expressed in Eq. (S41), x_P can be calculated using the peptide concentration and the degree of charge/protonation of all its units in the bulk solution.

- In the paper, we show the net charge of the peptide in the bulk solution:

$$Q_{pep} = \sum_\tau N_\tau f_\tau^b q_\tau \tag{S44}$$

Optimization with respect to f_{I^-} leads to

$$\frac{f_{I^-}}{1 - f_{I^-}} = \frac{K_I^0}{e^{\beta\mu_{H^+} - \beta\mu_{H^+}^0}} e^{-\beta\psi_S q_I} \tag{S45}$$

where K_I^0 is the thermodynamic equilibrium constant that describes the protonation/deprotonation of lipid head-groups, which satisfies

$$K_I^0 = e^{\beta\mu_{IH}^0 - \beta\mu_{I^-}^0 - \beta\mu_{H^+}^0} \tag{S46}$$

Using that $e^{\beta\mu_{H^+} - \beta\mu_{H^+}^0} = \dfrac{v_w \rho_{H^+}^b}{(v_w \rho_w^b)^{\frac{v_{H^+}}{v_w}}}$, we can re-express Eq. (S45) as

$$\frac{f_{I^-}}{1 - f_{I^-}} = \frac{(v_w \rho_w^b)^{\frac{v_{H^+}}{v_w}}}{v_w \rho_{H^+}^b} K_I^0 e^{-\beta\psi_S q_I} \tag{S47}$$

Optimization with respect to $\psi(z)$ leads to

the Poisson equation:

$$\begin{aligned} \frac{\partial^2 \psi}{\partial z^2} &= -\frac{\langle \rho_q(z) \rangle}{\epsilon_w} \; ; \quad z > 0 \text{ (in the solution)} \\ \frac{\partial^2 \psi}{\partial z^2} &= 0 \quad ; \quad -\frac{h_{mem}}{2} \leq z < 0 \text{ (inside the membrane)} \end{aligned} \tag{S48}$$

and the boundary conditions:

I. The electrostatic potential vanishes far from the membrane (in the bulk solution, $z \to \infty$):

$$\lim_{z \to \infty} \psi(z) = 0 \tag{S49}$$

II. In the membrane-solution interface ($z = 0$):

$$\epsilon_w \frac{\partial \psi(z)}{\partial z} \, |_{z=0, \text{ solution}} - \epsilon_M \frac{\partial \psi(z)}{\partial z} \, |_{z=0, \text{ membrane}} = \sigma_I f_{I^-} q_I \tag{S50}$$

III. Note that in writing Eq. (S48) we have imposed mirror symmetry of the system with respect to the plane $z = -\frac{h_{mem}}{2}$, which implies

$$\left. \frac{\partial \psi}{\partial z} \right|_{z=-h_{mem}/2} = 0 \tag{S51}$$

At this point the only remaining unknowns are the local interaction potentials $\pi_v(z)$ and $\psi(z)$ (including ψ_S). These local potentials are obtained after numerically solving the incompressibility constraint, Eq. (S25), the Poisson equation, Eq. (S48), for each z, and the boundary condition at the membrane-solution interface, Eq. (S50). Once these quantities have been calculated, all the functions that make the thermodynamic potential are determined and any equilibrium property can be derived.

Our theory is set in the semi-grand canonical ensemble, where the thermodynamic potential, whose minimum yields the equilibrium, is a function of the chemical potential

of all free species (see Eq. (S29)). The membrane composition, however, is fixed, and lipids can only modify the protonation state of their head-groups (f_{I^-}). In previous work, we have applied the same theory, but considered the case where lipids are allowed to leave or enter the system.[?] Within this mean-field treatment, changes in lipid composition are only important when there is no symmetry in the plane of the membrane surface, due to the presence of a pore for example (see Fig. 8 in **?**).

1.2 Bulk Solution

Let us consider the homogeneous solution that is in chemical equilibrium with the system described in Section 1.1. This bulk solution contains water molecules, hydronium ions, hydroxyde ions, salt ions, and peptides. Once, the pH, salt concentration, and peptide concentration are all set, the composition of this solution is completely defined. Namely, the concentration of each of the chemical species in the bulk solution is *experimentally* controlled.

The Helmholtz free energy of this bulk solution, F_b, can be expressed as

$$\begin{aligned}\beta\frac{F_b}{V} = &\sum_{\gamma\in\{w,H^+,OH^-,+,-\}} \rho_\gamma^b\Big(\ln\left(\rho_\gamma^b v_w\right) - 1 + \beta\mu_\gamma^0\Big)\\ &+ \sum_{\alpha_P\in\{\alpha_P\}} \rho_P^b(\alpha_P)\Big(\ln\left(\rho_P^b(\alpha_P) v_w\right) - 1 + \beta\mu_P^0\Big)\\ &+ \sum_{\tau\in\{\tau\}} \langle\rho_\tau^b\rangle\Big[g_\tau^b\big(\ln g_\tau^b + \beta\mu_{t,d}^0\big) + (1-g_\tau^b)\Big(\ln(1-g_\tau^b) + \beta\mu_{\tau,p}^0\Big)\Big]\end{aligned} \tag{S52}$$

where:

- V is the system's volume.
- All functions/quantities are independent of position. Superscript/subscript "b" indicates the value of the function in the bulk solution (sufficiently far from the

membrane). For example, the bulk degree of protonation of peptide τ units is

$$\lim_{z\to\infty} g_\tau(z) = g_\tau^b \tag{S53}$$

- We have arbitrarily chosen the value $\psi_b = 0$, for the constant electrostatic potential.
- The density of peptide's η units is

$$\langle \rho_\eta^b \rangle = N_\eta \sum_{\alpha_P \in \{\alpha_P\}} \rho_P^b(\alpha_P) = N_\eta \langle \rho_P^b \rangle \tag{S54}$$

which is true for all types of units, including titratable ones, $\{\tau\} \in \{\eta\}$.

In addition, the minimum of the thermodynamic potential must be consistent with two physical restrictions. First, the chemical solution is incompressible:

$$\sum_{\gamma\in\{w,H^+,OH^-,+,-\}} \rho_\gamma^b v_\gamma + v_P \langle \rho_P^b \rangle = 1 \tag{S55}$$

Second, the solution is charge-neutral:

$$\langle \rho_q^b \rangle = \sum_{\gamma\in\{H^+,OH^-,+,-\}} \rho_\gamma^b q_\gamma + \sum_{\tau\in\{\tau\}} \langle \rho_\tau^b \rangle f_\tau^b q_\tau = 0 \tag{S56}$$

where $\langle \rho_q^b \rangle$ is the density of electric charge.

The densities all free species in the bulk solution are a required input of any given calculation of the equilibrium conditions of the membrane-solution system, described in Section 1.1. The bulk peptide concentration defines $\langle \rho_P^b \rangle$ and $\langle \rho_\eta^b \rangle = N_\eta \langle \rho_P^b \rangle$ for each of its units. In addition, we assume that all n_c peptide conformations are equally probable in the bulk solution (see Eq. (S8)), which means that $\rho_P^b(\alpha_P) = \frac{\langle \rho_P^b \rangle}{n_c}$. The bulk solution pH determines $\rho_{H^+}^b$ and $\rho_{OH^-}^b$, using self-dissociation of water. Then, the salt concentration and the electroneutrality condition, Eq. (S56), are used to obtain ρ_+^b, ρ_-^b. Finally, ρ_w^b results

from solving the incompressibility of the bulk solution, Eq. (S55).

1.3 Numerical Method

To calculate results, we need to solve the incompressibility constraint, Eq. (S25), and the Poisson equation, Eq. (S48), with its boundary conditions. To such goal, we discretize the z-coordinate in layers of thickness $\delta = 0.5\,\text{nm}$. The number of layers considered is large enough so that both the local osmotic pressure, $\pi_v(z)$, and the electrostatic potential, $\psi(z)$, converge smoothly to their bulk values; typically we use $50 - 200$ layers (at 0.1 M salt concentration). To solve the equations we use an iterative Jacobian-free Newton-Krylov method and a FORTRAN 95 code developed in house. Typically, running an alpha-helix peptide from pH 1 to pH 10 with a step interval, $\Delta\text{pH} = 0.25$ can take around one week in a Intel Core i7-4790 CPU @ 3.60GHz. This time is greatly increased for the TAT peptide since we explicitly considered 5000 conformations of TAT. Simulation time depends mainly on the number of peptide conformations and the size of the peptide (and the salt concentration).

A repository with the code can be found at: https://github.com/BIOS-IMASL/patlm.

2 Comments on the Molecular Model

2.1 Value of χ_W

Tryptophan has various physical properties that makes it the optimal amphipathic residue for the stabilization of peptides and proteins at water-lipid membrane interfaces.[???] To model the adsorption of hydrophobic residues to lipid bilayers, we have included a term in the thermodynamic potential that describes the internal energy of tryptophan-

membrane contacts, U_{AD}. This adsorption energy is quantified by the Flory-Huggins interaction parameter χ_W. The particular value assigned to the tryptophan-membrane interaction parameter deserves a special note.

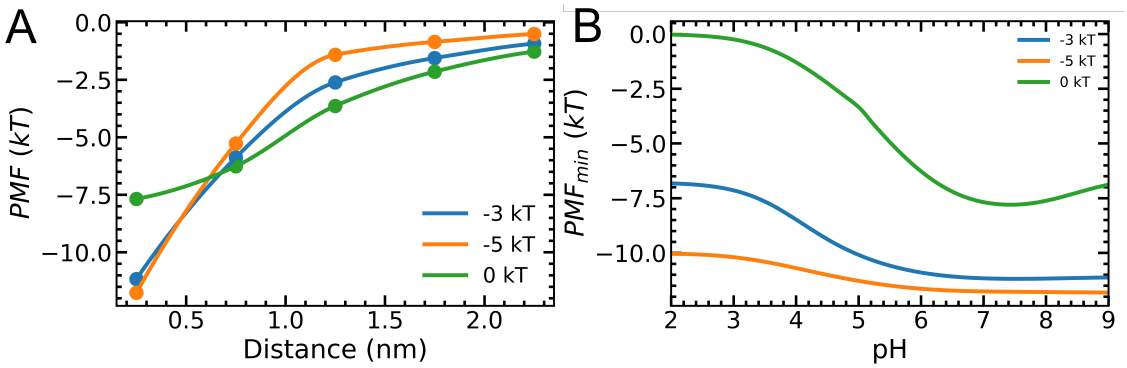

Figure S1: (A) Local PMF and (B) PMF_{min} of R6W3 for different tryptophan χ parameters. [CPP] $= 1\,\mu$M, [NaCl] $= 0.1$ M and pH 7; membrane has 20% ionizable lipid with pKa $= 4.5$. PMF_{min} corresponds to the global minimum of $PMF(z)$ (z =distance from membrane surface).

Dimerization free energies of small hydrophobic molecules at room temperature are generally around $-10\,\text{kJ mol}^{-1}$ ($-4k_BT$).[?] For example, the vapor pressure studies of **?** reported free energy values of -2 to $-2.7\,\text{kcal mol}^{-1}$ (-3.4 to $-4.6k_BT$) for the dimerization of some hydrocarbons in water near room temperature. Using different spectroscopy techniques, **?** suggested that the molecular interactions between tryptophan and a lipid bilayer at the interface with water can aid and stabilize protein folding by approximately $2\,\text{kcal mol}^{-1}$ ($3.4k_BT$).

Then, we use $\chi_W = -3k_BT$ in this work, but in order to assess the effect of this parameter, we have performed calculations using $\chi_W = 0$ and $-5k_BT$ as well. Figure S1 shows results for the adsorption of R6W3 under identical conditions but using these different χ_W values. When $\chi_W = 0$ at very acidic pHs, $PMF_{min} = 0$. This is because at this pH the lipids are fully protonated, which results in no electrostatic interactions with the peptide; there is no driving force for peptide adsorption. When $\chi_W < 0$ there is an attractive PMF_{min} at acidic pHs, which becomes more negative when χ_W decreases from -3 to

$-5k_BT$.

2.2 Peptide Configurations

Random coil peptides are modeled using a Rotational Isomeric States (RIS) method,[?] in which each 0.38 nm-long segment can assume one of three isoenergetic orientations.[?] The number of generated conformations n_c using the RIS model is in the order of $n_c = 3^{L-2}$; where L is the number of residues in the sequence. In this way, n_c grows exponentially with L. For $L > 9$ the total number of conformations increases the computational cost of the simulations sufficiently to render our calculations impractical. Particularly problematic is the TAT peptide having $L = 13$. To solve the intractability of this peptide, after each of the generated conformations were rotated 50 times using randomly generated Euler angles, 5000 of these conformations were randomly selected to be included in the calculations. To maintain the same ratio of selected/total conformations for both random coil peptides, 69 previously-rotated conformations of the R9 peptide are randomly included in our calculations.

3 Additional Results

3.1 Adsorption Isotherms

The adsorption isotherms for the different peptides are displayed in Figure S2A. Panel B shows PMF_{min}, the global minimum of $PMF(z)$, for the same conditions; PMF_{min} represents the adsorption free energy.

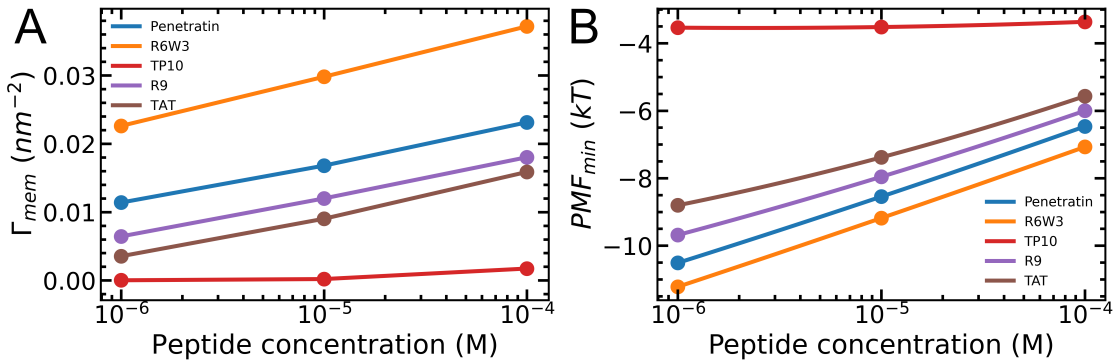

Figure S2: (A) Adsorption excess, Γ_{mem}, and (B) adsorption free energy, PMF_{min}, as a function of peptide concentration (in the bulk solution); pH 7 and [NaCl] = 0.1 M; the membrane has a 20% of ionizable lipid with pKa = 4.5.

3.2 Ionizable Lipid Fraction

Additional results on the effects of the area fraction occupied by ionizable lipid are presented in Figure S3. Panel A shows the global minimum of the $PMF(z)$, PMF_{min} as function of the solution pH for R9 solutions (where z = is the distance from the membrane surface). The different curves correspond to membranes having different acidic lipid composition: 10%, 20% and 40% of the total area. Panel B of Figure S3 presents the adsorption, Γ_{mem}, at pH 7 as function of the fraction of ionizable lipid in the membrane.

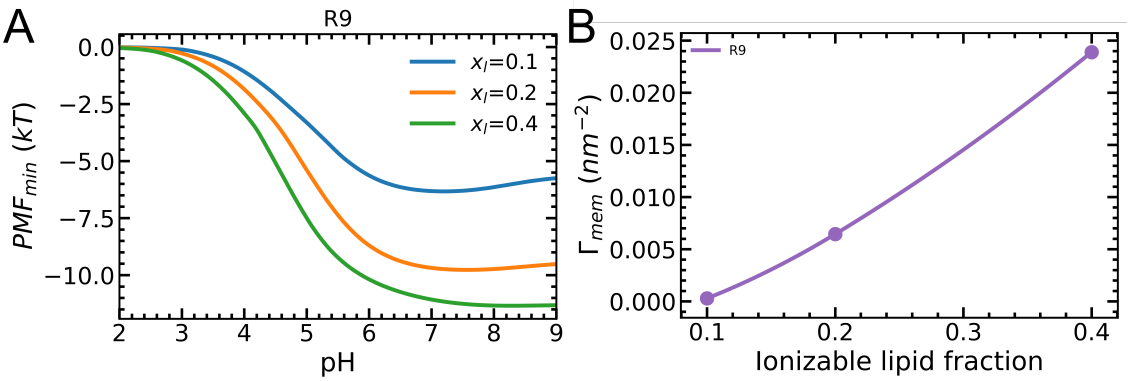

Figure S3: A: PMF minimum as a function of pH for membranes having different area fraction of ionizable lipids, x_l. B: Adsorption as a function of the fraction of ionizable lipids at pH 7. In both panels [CPP] = 1 μM and [NaCl] = 0.1 M; lipid pKa = 4.5.

3.3 Membrane Surface Charge and Local pH

In Figure S4 we show that CPP adsorption (at [CPP] = 1 μM) does not significantly affect the charge density on the membrane surface, σ_{I^-}. Due to dissociation of the ionizable lipids' head-groups, surface charge depends strongly on the solution pH.

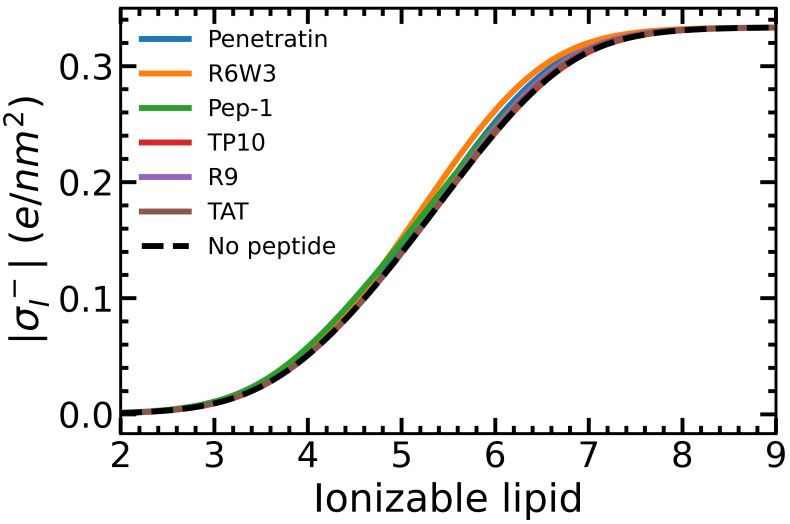

Figure S4: Surface charge density on the membrane surface (in absolute value) as a function of pH for different 1 μM peptide solutions. The dashed line corresponds to a solution without peptides ([CPP] = 0). [NaCl] = 0.1 M; membrane has a 20% of ionizable lipid with pKa = 4.5.

The small changes in charge density following the adsorption of different peptides, seen in Figure S4, result from lipid protonation/deprotonation; upon peptide adsorption, a different local pH establishes near the membrane surface. This behavior is illustrated in Figure S5 that displays local pH profiles for the adsorption of different CPPs. In addition, Fig. S6 shows local pH profiles for the different permutations of R6W3 discussed in Figs. 6 and 7 of the paper.

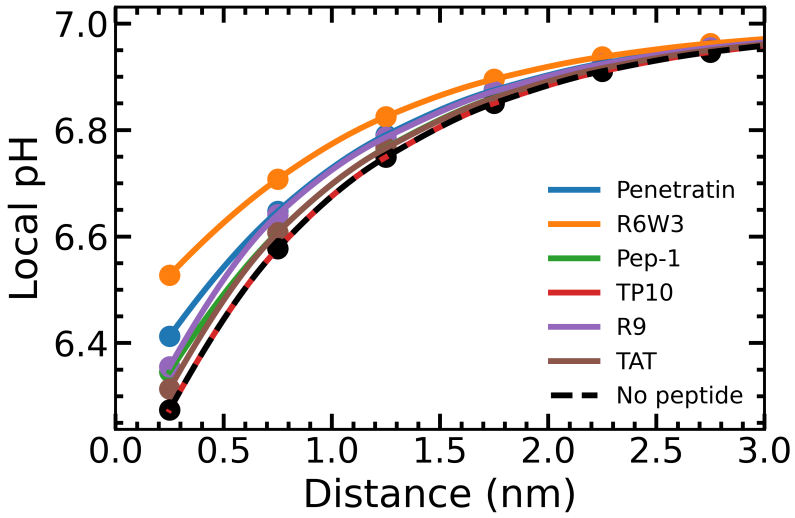

Figure S5: Local pH as a function of the distance from the membrane, for different peptide solutions; [NaCl] = 0.1 M, [CPP] = 1 μM (solid lines) and pH 7; membrane has a 20% of ionizable lipid with pKa = 4.5. These conditions correspond to those of Fig. 3 in the paper.

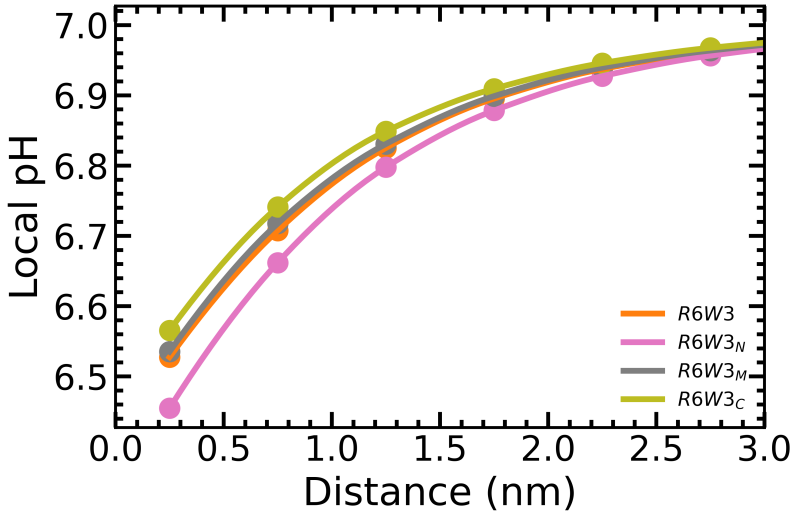

Figure S6: Local pH as a function of the distance from the membrane, for different R6W3 permutations; [NaCl] = 0.1 M, [CPP] = 1 μM (solid lines) and pH 7; membrane has a 20% of ionizable lipid with pKa = 4.5. These conditions correspond to those of Fig. 7 in the paper.

3.4 Local Water and Salt Ions Distributions

How are water molecules and salt ions distributed near the membrane surface? To answer this question, Fig. S7 shows the volume fraction of water as a function of the distance

to the membrane surface. Similarly, Fig. S8 shows the concentration of salt anions as a function of position, while Fig. S9 shows that of salt cations. These figures include curves for different peptides solutions (solid lines), which correspond to the same conditions of Fig. 3 of the paper; distributions corresponding to solutions without peptides are also included in these graphs (dashed lines).

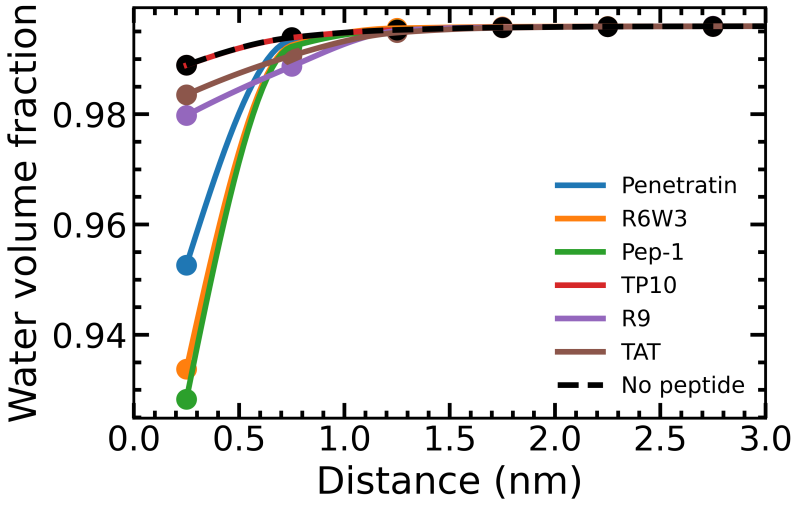

Figure S7: Local water volume fraction as a function of the distance from the membrane, for different peptide solutions; [NaCl] = 0.1 M, [CPP] = 1 μM (solid lines) and pH 7; membrane has a 20% of ionizable lipid with pKa = 4.5. These conditions correspond to those of Fig. 3 in the paper.

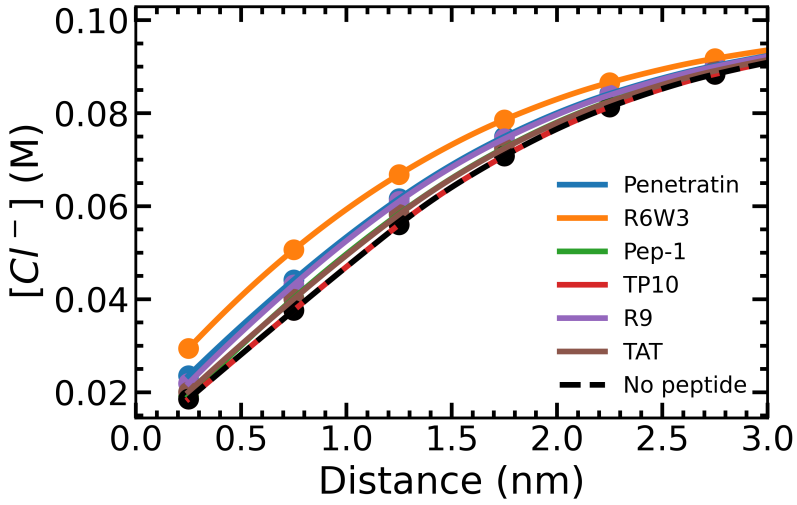

Figure S8: Local chloride ion concentration as a function of the distance from the membrane, for different peptide solutions; [NaCl] = 0.1 M, [CPP] = 1 μM (solid lines) and pH 7; membrane has a 20% of ionizable lipid with pKa = 4.5. These conditions correspond to those of Fig. 3 in the paper.

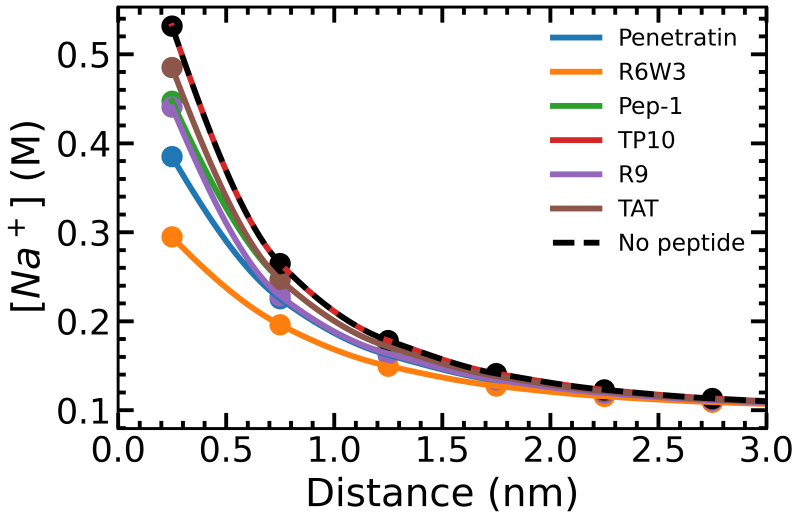

Figure S9: Local sodium ion concentration as a function of the distance from the membrane, for different peptide solutions; [NaCl] = 0.1 M, [CPP] = 1 μM (solid lines) and pH 7; membrane has a 20% of ionizable lipid with pKa = 4.5. These conditions correspond to those of Fig. 3 in the paper.

3.5 Volume Fraction Distributions: TAT and Penetratin

Figure 5 in the paper shows the volume fraction occupied by different groups of amino-acids as a function of the distance from the membrane surface for some CPPs. Figure S10

shows those volume fractions profiles for TAT and Penetratin.

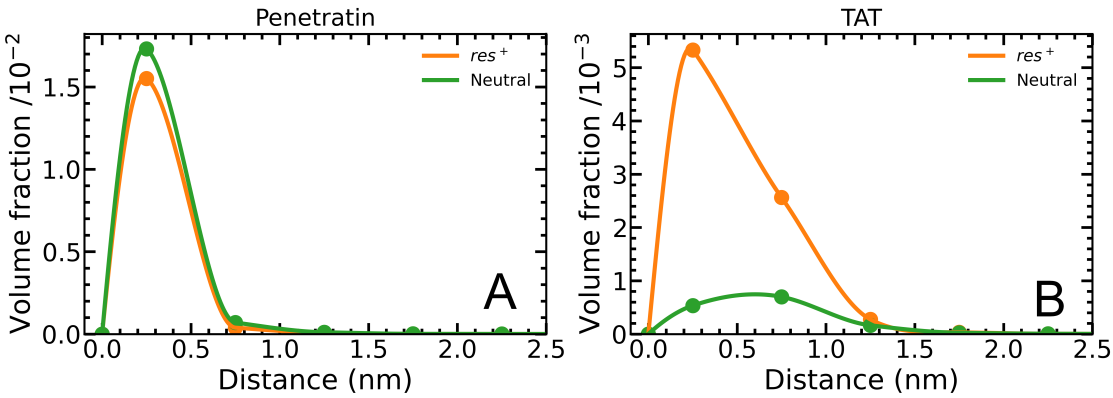

Figure S10: Local volume fractions occupied by each type of residue as a function of the distance to the membrane surface ($z = 0$) for Penetratin (A) and TAT (B) peptides. Residues are divided into positively (res^+) and negatively charged (res^-), and charge neutral, according to their charge at pH 7. The membrane has 20% ionizable lipid with pKa $= 4.5$; [CPP] $= 1\,\mu$M, [NaCl] $= 0.1\,$M and pH 7. Solid lines correspond to point interpolation using Akima splines.