2022

Free-energy landscapes and insertion pathways for peptides in membrane environment

Ganga P. Sharma

Aaron C. Meyer

Suhail Habeeb

Michael Karbach

Gerhard Müller

Free-energy landscapes and insertion pathways for peptides in membrane environment

Ganga P. Sharma,[1] Aaron C. Meyer,[1] Suhail Habeeb,[1] Michael Karbach,[2] and Gerhard Müller[1]

[1] *Department of Physics, University of Rhode Island, Kingston RI 02881, USA*

[2] *Fachgruppe Physik, Bergische Universität Wuppertal, D-42097 Wuppertal, Germany*

Free-energy landscapes for short peptides – specifically for variants of the pH Low Insertion Peptide (pHLIP) – in the heterogeneous environment of a lipid bilayer or cell membrane are constructed, taking into account a set of dominant interactions and the conformational preferences of the peptide backbone. Our methodology interprets broken internal H-bonds along the backbone of a polypeptide as statistically interacting quasiparticles, activated from the helix reference state. The favored conformation depends on the local environment (ranging from polar to nonpolar), specifically on the availability of external H-bonds (with H_2O molecules or lipid headgroups) to replace internal H-bonds. The dominant side-chain contribution is accounted for by residue-specific transfer free energies between polar and nonpolar environments. The free-energy landscape is sensitive to the level of pH in the aqueous environment surrounding the membrane. For high pH, we identify pathways of descending free energy that suggest a coexistence of membrane-adsorbed peptides with peptides in solution. A drop in pH raises the degree of protonation of negatively charged residues and thus increases the hydrophobicity of peptide segments near the C terminus. For low pH, we identify insertion pathways between the membrane-adsorbed state and a stable trans-membrane state with the C terminus having crossed the membrane.

I. INTRODUCTION

Water-soluble peptides with an affinity to insertion into cell membranes under specific conditions have found applications in medical research as diagnostic and therapeutic tools. They have been shown to carry fluorescent markers or drugs to targeted regions in living organisms with a continually advancing degree of efficiency and discrimination [1–9]. Yet many open questions remain and call for answers. Analytic studies and computational studies tend to be complementary in their strengths and limitations to provide answers. Both approaches are needed to advance our knowledge of peptide insertion into lipid membranes. Molecular dynamics studies offer high resolution at short time scales, whereas kinetic studies, which combine analytic and computational aspects in different ways, are better equipped for dealing with processes that involve multiple time scales.

This work reports the second stage of a three-stage project aiming to support a more complete theoretical understanding of membrane-associated protein or peptide folding [10–23]. The first stage, reported in Ref. [24], involved the design and solution of a microscopic model for the coil-helix transition of a long polypeptide adsorbed to a planar water-lipid interface. The methodology interprets the polypeptide and its homogeneous, effectively two-dimensional environment as a system of statistically interacting quasiparticles activated from the (ordered) helix state. The particles represent broken internal H-bonds along the backbone, producing segments of a (disordered) coil conformation, sprawled across the interface in the shape of a self-avoiding random walk.

In one experimentally realized scenario [2, 9, 25–35], the coil-helix conversion is triggered by a drop in pH. The ensuing protonation of negatively charged side chains enhances the hydrophobicity of the polypeptide and pushes its backbone deeper into the (nonpolar) membrane. This environmental change favors the formation of internal hydrogen bonds, which stabilize the α-helix conformation. Depending on the parameter settings, the model predicts the conformation to change as a crossover, a first-order transition, or a second-order transition.

Here, in the continuation of this project, we begin by considering long polypeptides which are no longer confined to a flat water-lipid interface, but are positioned and oriented along some line in the heterogeneous environment comprising the lipid bilayer of a liposome or a biological cell and the surrounding water. The model parameter identified in Ref. [24] to drive the conformational change then turns into a (scalar) field with values reflecting the nature of the local medium, specifically its degree of polarity. Such circumstances pose a challenge to any method of analysis. The methodology used here has already proven to be adaptable to heterogeneous environments in different applications [36–38].

From the analysis thus carried out, profiles emerge for the densities of free energy, enthalpy, entropy, and helicity pertaining to segments of long polypeptides in the heterogeneous membrane environment. The profiles reflect enthalpic and entropic consequences of the interactions between the backbone of the polypeptide and the lipid membrane or the surrounding hydrogen-bonded network of water molecules. These profiles are then taken to represent propensities for the statistical mechanical behavior of short peptides in the same environment. In this step, enthalpic and entropic effects involving the side chains and the semi-fluid bilayer of lipid amphiphiles can be built into the model. The result are landscapes of free energy, enthalpy, entropy, and helicity for short peptides of given composition.

The free-energy landscapes incorporate road signs for pathways that guide the peptide from solution to adsorption under one set of environmental circumstances and from adsorption to insertion under a different set. The

enthalpy and entropy landscapes offer insights into the dominant forces that shape the pathways of descending free energy. This analysis, which is grounded in equilibrium statistical mechanics, will be a key ingredient for the third stage of this project [39]: a kinetics study of trans-membrane insertion and exit of peptides.

The theoretical study reported here is custom-made for variants of the *pH Low Insertion Peptide* (pHLIP) family [28–30, 32, 40, 41], but not to the exclusion of other peptides with similar attributes. The sequences of three pHLIP variants are shown in Fig. 1 with some relevant features highlighted. The positioning in the sequence of charged residues and polar residues is instrumental for the solubility of pHLIP in water. The hydrophobic residues provide the affinity for adsorption of pHLIP to the water-lipid interface. The protonatable negative charges at and near the (inserting) C terminus make pHLIP sensitive to the experimentally controllable and reversible change in pH.

FIG. 1: Sequence of the $N_R = 32$ residues that make up variants W6, W17, W30 (top to bottom) of pHLIP. Highlighted are the residues Arg, Asp, which have electrically chargeable side chains, and the Trp residue used as a marker in fluorescent spectroscopy.

The Trp residues are markers for fluorescent spectroscopy, by which the insertion and exit processes are monitored [30, 42, 43]. The three pHLIP variants with Trp at different positions in the sequence of residues are designed to yield clues about the modes of insertion and exit. The conformational changes between coil and helix segments that accompany both insertion and exit can be monitored by circular dichroism (CD) spectroscopy [30, 31]. The helix-inhibiting Pro residue near the center of the sequence may be instrumental for the mechanics of the insertion process, allowing a kink between two fully formed helical segments [44–47]. Peptide insertion into membranes is, of course, a wider field of research beyond the limited focus of this study. There are aggregates of peptides which interact with membranes in significantly different ways, by forming pores, for example. In this work no peptide aggregates are being considered.

The following sections are about fields, profiles, landscapes, and pathways: fields of environmental parameters (Sec. II), profiles for local properties of long polypeptides (Sec. III), landscapes for global properties of short peptides (Sec. IV), and pathways associated with descending free energy of short peptides (Sec. V). Finally, we briefly discuss effects not yet accounted for and outline the challenges facing kinetic studies (Sec. VI).

II. MEMBRANE ENVIRONMENT

All heterogeneity in a lipid bilayer considered here is associated with the normal spatial coordinate x. We set $x = 0$ at the center of the bilayer. The outside of the cell or liposome is at positive x. Any effects of curvature are set aside as higher-order corrections to results presented here. The membrane environment is characterized by several parameters. We take the dominant parameter field to be the concentration of H_2O molecules. Hydrophobic interactions are prevalent [48]. Subdominant parameter fields involve electrostatic interactions including trans-membrane, surface, and dipole potentials [49]. Further parameters are related to properties of lipids, notably the profile of lateral pressure and the entropy reduction along the contact line with the peptide [40, 50–53]. We examine the effects of the dominant environmental parameter in some detail and discuss those of subdominant parameters summarily at the end.

A. Density field of water

The dominant environmental parameter, the density field $\rho_w(x)$ of H_2O molecules, is symmetric under reflection about $x = 0$. It is a dimensionless quantity varying between $\rho_w(x) = 1$ sufficiently far from the lipids and $\rho_w(x) \ll 1$ near the center of the bilayer. We use a smoothed-ramp density field as a model representation in our statistical mechanical analysis:

$$\rho_w(x) = 1 - \frac{x_s}{x_a - x_b} \ln\left(\frac{\cosh\left(\frac{x}{x_s}\right) + \cosh\left(\frac{x_a}{x_s}\right)}{\cosh\left(\frac{x}{x_s}\right) + \cosh\left(\frac{x_b}{x_s}\right)}\right). \quad (1)$$

It has two control parameters, $x_b/x_a > 1$ and $x_s/x_a > 0$. A density field of such shape (Fig. 2) is well-established from experiment [54] and computer simulations [55].
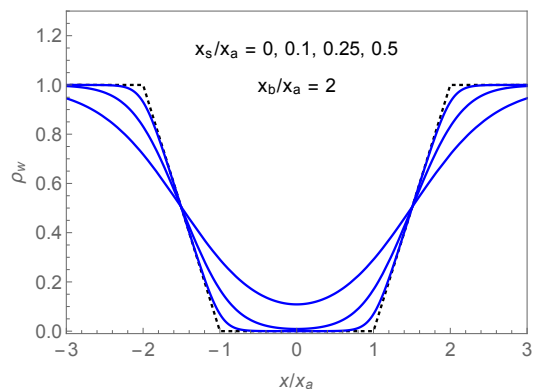
FIG. 2: Smoothed-ramp density field of H_2O molecules. The parameters x_a and x_b in (1) locate the bottom and the top of the ramp symmetrically, and x_s controls the softening of the edges.

Throughout Secs. II-IV we use the following specifications: $x_a = 15$Å represents the distance from the center of the bilayer to a point just inside the lipid headgroups, and $x_b = 25$Å the distance from the center to a point outside, where H_2O molecules are in contact with the polar ends of the lipid headgroups; the value $x_s \simeq 3$Å ensures smoothness over an atomic length scale [56]. The results presented in the following are not sensitive to small variations in these parameters.

A peptide of 32 residues in helix conformation would have a length of roughly 50Å, if we assume that each helical link [57] adds an element $l_h \simeq 1.5$Å of length in the direction of the axis [58]. The length (end-to-end distance) of the peptide in the coil conformation is a fluctuating quantity. With a size $l_a \simeq 4$Å for each link, the contour length of the peptide becomes more than double its length in helix conformation. The coil conformation, modeled as a random walk, has an average end-to-end distance of ~ 24Å if unrestricted. Geometrical and dynamical constraints make the average end-to-end distance considerably longer [59].

B. Free-energy

The term *free-energy landscape* of the title requires some explanation. The system under consideration includes a peptide in an environment consisting of a lipid bilayer surrounded by water. For a homogeneous system, the Gibbs free energy can be expressed in the form,

$$G(T, P) = U + PV - TS = H - TS, \tag{2}$$

where U is the internal energy, H the enthalpy, S the entropy, and V the volume. The pressure P and the temperature T are control variables.

Our system is not homogeneous in all respects. We consider situations at uniform T, typically room temperature ($T_{rm} = 293$K). The pressure is uniform in the aqueous environment and the normal pressure, P_N, also across the bilayer. However, the lateral pressure, P_L, has a characteristic profile that averages out to the value of the normal pressure [60–69]. For the purpose of this study, we only consider quasistatic processes of a restricted type in which both T, and P_N remain constant and uniform. One natural energy scale uses $k_B T_{rm} \simeq 4.0 \times 10^{-21}$J $\simeq 0.58$kcal/mol as its unit. The processes involve the translocation of the peptide, accompanied by changes in conformation of the peptide and in its interactions with water and lipids.

Each quasistatic process can be described as a path in a space of independent variables. The variation of G ascends or descends in a landscape of sorts, just as a path on a topographic map does. Path segments with $\Delta G < 0$ are favorable and path segments with $\Delta G > 0$ unfavorable regarding spontaeneous occurrence. Paths that are all downhill are likely to have fast realizations in experiments. By contrast, paths that have significant barriers between initial and final points are much slower if realized at all.

All changes of G are a combination of enthalpic and entropic contributions. For the processes under consideration here we can write,

$$\Delta G = \Delta H - T\Delta S. \tag{3}$$

We refer to changes $\Delta H < 0$ as associated with an *enthalpic gain* and changes $\Delta S > 0$ as associated with an *entropic gain* in the sense that a gain lowers G while a loss does the opposite.

Entropic losses, $\Delta S < 0$, that are relevant for this study include the following: (i) a complete or partial conformational change of the peptide from (disordered) coil to (ordered) helix; (ii) the immobilization of H_2O molecules via the formation of H-bonds with polar contacts on the backbone or the side chains of the peptide or with polar contacts on lipid head-groups; (iii) the formation of an ordered contact line between lipid head-groups and the peptide in adsorbed or trans-membrane states.

Enthalpic losses related to a change $\Delta U > 0$ are all associated with molecular interactions. The two main sources of this type in the context of this study involve (i) the breaking of H-bonds and (ii) the translocation of charges or polar contacts from a polar environment (water) into a nonpolar environment (membrane). The H-bonds in question include internal ones between backbone contacts, and external H-bonds between the peptide (backbone or side chain), H_2O molecules, and lipid head-groups, all of which have polar contacts. A different type of enthalpic loss, $(P_L^h - P_L^l)\Delta V > 0$, comes into play when a peptide segment (e.g. a residue) of volume ΔV translocates from a position of low lateral pressure, P_L^l, to a position of high lateral pressure, P_L^h.

It is quite challenging, in general, to estimate all these contributions with some accuracy. Existing estimates found in the literature vary widely, in part due to differences in underlying assumptions. In what follows, a case will be made for identifying and quantifying dominant contributions.

C. Enthalpic cost of H-bonds

In the α-helix conformation the backbone of each residue participates in two H-bonds. The CO group of residue n is acceptor to the NH group of residue $n + 4$ acting as donor. The helix conformation thus involves one internal H-bond per residue. The conversion of a helix segment into a coil segment breaks internal H-bonds, for which there is an enthalpic cost.

Deep inside the lipids the full price is due, up to 5kcal/mol per H-bond, which is considerable in units of ambient thermal fluctuations. When the peptide is positioned in the polar environment of the lipid headgroups and the adjacent water, there are opportunities for internal H-bonds to be replaced by external ones. The

replacement bonds reduce the (enthalpic) cost of breaking internal H-bonds. At the same time, it increases the maximum number of H-bonds per residue from one to two. The replacement bonds (with reduced directionality) are likely to be weaker.

Whether the enthalpic bottom line in this case is a gain or loss depends on how the exposed backbone of the polypeptide affects the internal H-bonds of liquid water. In ice there are two intact H-bonds per H_2O molecule. Each molecule shares four bonds, two in the role of donor and two as acceptor. In liquid water about 12% of H-bonds between H_2O molecules are broken [58, 70]. The intact H-bonds form a dynamic network with H_2O molecules, sharing less than four bonds on average.

If the fraction of unsatisfied H-bonds between H_2O molecules is higher at the lipid-water interface than inside the bulk, then the exposed CO groups and NH groups along a coil segment of the backbone are more likely to encounter partners for external H-bonds in adjacent H_2O molecules. This reduces the need for breaking H-bonds between H_2O molecules when external H-bonds are formed along the backbone of the peptide. The formation of external H-bonds in water is then more likely to result in an enthalpic gain.

The methodology developed in Ref. [24] encodes the enthalpic contribution to the conformational affinity in the activation energy ϵ of a coil link from the helix reference state. The heterogeneous membrane environment is accounted for by turning this parameter into a field. We use the ansatz,

$$\epsilon(x) = \epsilon_{\mathrm{Hb}}\big[1 - \alpha_{\mathrm{H}}\rho_{\mathrm{w}}(x)\big], \tag{4}$$

for the dependence on position, the premise being that the density field $\rho_{\mathrm{w}}(x)$ of water represents the dominant environmental influence. Near the center of the lipid bilayer we have $\rho_{\mathrm{w}}(x) \ll 1$, which maximizes $\epsilon(x)$ to roughly the strength of an internal H-bond. At positions closer to the lipid-water interface, $\epsilon(x)$ decreases as $\rho_{\mathrm{w}}(x)$ increases. This change is due to the growing probability that internal H-bonds are being replaced by external ones.

The parameter α_{H} determines whether in the aqueous environment, with $\rho_{\mathrm{w}}(x) \simeq 1$, we have an enthalpic gain ($\alpha_{\mathrm{H}} > 1$) or an enthalpic loss ($\alpha_{\mathrm{H}} < 1$). Here we use $\epsilon_{\mathrm{Hb}}/k_BT \simeq 9$, i.e. $\epsilon_{\mathrm{Hb}} \simeq 5\mathrm{kcal/mol}$. The strength of H-bonds varies and depends on geometrical constraints and the charge distribution on the polar contacts. H-bonds between H_2O molecules in liquid water are highly fluctuating in strength, and could be as low as $2\mathrm{kcal/mol}$ on average [71–76].

D. Entropic cost of H-bonds

The enthalpic cost reduction associated with external H-bonds in the polar environment of liquid water comes with an entropic price that has yet to be included in the accounting. Every H-bond formed between an exposed backbone CO or NH group with an H_2O molecule reduces the mobility of that water molecule to some extent and thus lowers its contribution to the entropy.

It is hard to estimate the magnitude $|\Delta \bar{S}_{\mathrm{H}}|/k_{\mathrm{B}}$ of this entropy reduction from first principles. At a glance, we might expect it to be comparable in magnitude to the entropic gain per residue produced when a segment of (ordered) helix transforms into a segment of (disordered) coil. In Ref. [24] we have calculated that entropy gain per residue to range between $\ln 2 \simeq 0.7$ and $\ln 3 \simeq 1.1$ in units of k_{B}. Backbone segments in coil conformation are less mobile than H_2O molecules in the dynamic network of H-bonds that make up liquid water. The unstructured nature of the coil reduces the entropy loss of H_2O molecules forming H-bonds with it. All this makes it reasonable to operate with an estimated upper limit,

$$\frac{|\Delta \bar{S}_{\mathrm{H}}|}{k_{\mathrm{B}}} \lesssim 1. \tag{5}$$

The actual value of $|\Delta \bar{S}_{\mathrm{H}}|$ could be lower on account of the role played by the lipid headgroups. The zwitterionic headgroups of POPC, for example, offer alternative contacts for external H-bonds to backbone segments in coil conformation. They come at lower entropic cost.

The overall message for what follows in Sec. IV is that the entropic contribution to the free energy of the peptide backbone in coil conformation is significantly reduced from what the disorder of the coil conformation would suggest when the peptide is in contact with water.

E. Model for peptide conformation

In Ref. [24], we exactly solved a microscopic statistical mechanical model for the conformational transformation between coil and helix of a long polypeptide positioned in a plane parallel to the interface of a polar and a nonpolar medium such as realized by water and a lipid bilayer. Depending on the settings and variations of its parameters, the model predicts a conformational crossover or a transition of first or second order between coil and helix.

All microstates of the peptide are characterized, in this model, by combinations of 2μ species of statistically interacting nested particles: hosts ($m = 1$) nucleate coil segments, whereas hybrids ($m = 2, \ldots, \mu$) and tags ($m = \mu + 1, \ldots, 2\mu$) grow such segments in two perpendicular directions of a self-avoiding random walk as illustrated in Fig. 3.

The activation energies ϵ_m, $m = 1, \ldots, 2\mu$, can be tailored to meet the physical requirements at hand. In the plane of the water-lipid interface all activation energies are spatially uniform. Nucleating a coil segment requires that several internal H-bonds along the backbone of the α-helix are broken simultaneously, whereas the growth of a coil segment already nucleated proceeds by the sequential breaking of single H-bonds.

The model, therefore, assigns different activation energies for the control of nucleation and for the control of
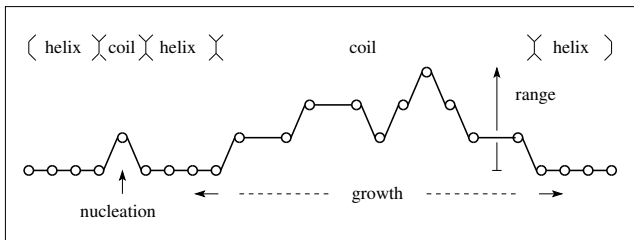
FIG. 3: Coil conformation of polypeptide modeled as a self-avoiding random walk generated by statistically interacting nested particles activated from the helix reference state via *nucleation* (control parameter τ) and *growth* (control parameter t) with *range* limited by control parameter μ.

growth, namely $\epsilon_1 \doteq \epsilon_n$ for hosts, $\epsilon_2 = \cdots = \epsilon_\mu \doteq 2\epsilon_g$ for hybrids, and $\epsilon_{\mu+1} = \cdots = \epsilon_{2\mu} \doteq \epsilon_g$ for tags. These activation energies in units of the thermal energy k_BT are usefully expressed by the nucleation parameter τ (also named cooperativity) and growth parameter t:

$$\tau = e^{\beta(\epsilon_g - \epsilon_n)} \quad : 0 < \tau \leq 1, \tag{6a}$$

$$t = e^{\beta\epsilon_g} \quad\quad : 0 < t < \infty. \tag{6b}$$

In addition to these two continuous control parameters, the discrete model parameter μ controls the range of the random walk away from the axis of the local helix segments. All model features from Ref. [24] used in the following are reviewed in Appendix A.

F. Model parameter field

Of the three model parameters, the growth parameter t is the one most sensitive to the environment by far. We begin the model adaptation to the membrane environment by keeping the cooperativity parameter τ and the range parameter μ uniform, while we turn t into a field. For this purpose, we use the ansatz (4) as we construct two fields of scaled activation energy, $\epsilon_n(x)$ for hosts, $\epsilon_g(x)$ for hybrids and tags [77]. We leave the former ($m = 1$) undetermined for now and link the latter ($m = 2, 3, \ldots, 2\mu$) to the density field of water:

$$K_n(x) \doteq \frac{\epsilon_n(x)}{k_BT}, \tag{7}$$

$$K_t(x) \doteq \frac{\epsilon_g(x)}{k_BT} = \frac{\epsilon_{Hb}}{k_BT}\left[1 - \alpha_H\rho_w(x)\right], \tag{8}$$

where ϵ_{Hb}/k_BT represents the scaled energy of an H-bond and $\alpha_H \simeq 1$ is the enthalpy parameter introduced previously (Sec. II C), assumed to be equal for hybrids and tags. The growth parameter field,

$$t(x) = e^{K_t(x)} \quad : 0 < t(x) < \infty, \tag{9}$$

is environmentally sensitive via the shape of $\rho_w(x)$ and the value of α_H. The cooperativity,

$$\tau = e^{K_t(x) - K_n(x)} = t(x)e^{-K_n(x)} \quad : 0 < \tau \leq 1, \tag{10}$$

is kept as a position-dependent parameter. The function $K_n(x)$ is determined from (10).

This choice of modeling is informed by the following facts. Cooperativity, which controls the nucleation of coil segments, is a process initiated by thermal fluctuations within the backbone of an intact segment of α-helix. Multiple H-bonds must be broken simultaneously. They are all protected from direct environmental contact. Nucleation is only affected indirectly by an environmental change from nonpolar to polar. In the nonpolar environment, the nucleation energy barrier is followed by a high plateau and in the polar environment by a low plateau. The former favors a reversal of nucleation events, whereas the latter favors the growth of nucleated segments.

Figure 4(a) shows the model density field of water $\rho_w(x)$ used henceforth. It has the smoothed-ramp profile (1). The growth parameter field $t(x)$ inferred from the predominant environmental field $\rho_w(x)$ via (8) and (9) is shown in Fig. 4(b). When the polypeptide is in coil conformation while adsorbed to the water-lipid interface, its position is near the outer dot-dashed line. For the coil conformation to be stable, the growth parameter $t(x)$ must be below the upper horizontal dashed line (at $t = 3$). That is barely the case if $\alpha_H < 1$. A robust coil conformation requires that $\alpha_H \gtrsim 1$, implying that breaking internal H-bonds along the backbone of the polypeptide and replacing them by external H-bonds with available H_2O molecules is an enthalpic gain.
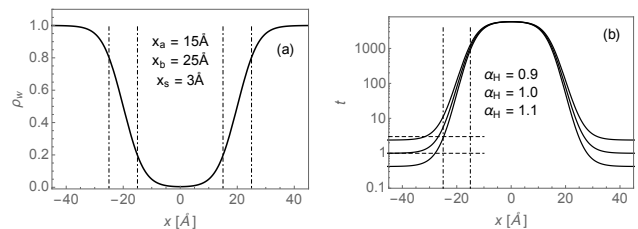

FIG. 4: (a) Model density field of water $\rho_w(x)$ and (b) growth parameter field $t(x)$, with specifications as indicated. The dot-dashed lines represent the endpoints of the ramp and the dashed lines the range $1 \leq t \leq 3$ in which a coil-helix phase transition may occur.

The extended model system is now ready for applications to the heterogeneous membrane environment. We have converted t into the field (9) and kept the control parameters τ and μ uniform. We have already stated reasons for not turning the nucleation parameter τ into a field. Regarding the discrete parameter μ [78], we will consider the cases $\mu = 2$ and $\mu = \infty$, for which analytic solutions are available in Ref. [24]. The two values span a range that is correlated with a range of entropy generated inside coil segments of given length.

III. PROFILES

The analysis reported in the following yields profiles for specific position-dependent attributes of a polypeptide backbone in the heterogeneous membrane environment. The relevant quantities are the helicity $\bar{N}_{\mathrm{hl}}$ as well as densities of free energy $\bar{G}$, enthalpy $\bar{H}$, and entropy $\bar{S}$. What is taken into account at this stage are the internal H-bonds along the backbone of a long, generic polypeptide and external H-bonds with water or lipid headgroups depending on their local availability. Also taken into account is the entropy of the backbone in coil conformation and an entropic contribution associated with external H-bonds (Sec. II D).

A. Internal H-bonds

We construct profiles from the expressions for helicity $\bar{N}_{\mathrm{hl}}(t,\tau)$, entropy density $\bar{S}(t,\tau)$, free-energy density $\bar{G}(t,\tau)$, and enthalpy density $\bar{H}(t,\tau)$ stated in Appendix A, in combination with the field $t = t(x)$ for the growth parameter (9), and a value of choice for the nucleation parameter τ, Eq. (10). The growth parameter field $t(x)$ determines all profiles via local relations. Profiles for $\mu = 2$ (narrow range of conformational disorder) are shown Fig. 5 and profiles for $\mu = \infty$ (broad range) in Fig. 6.

Well inside the lipid bilayer, at $x \simeq 0$, the helix conformation is firmly established. All internal H-bonds are intact. There is no configurational disorder. Therefore, the order parameter (helicity) is close to saturation whereas the densities of enthalpy [79] and entropy are near zero.
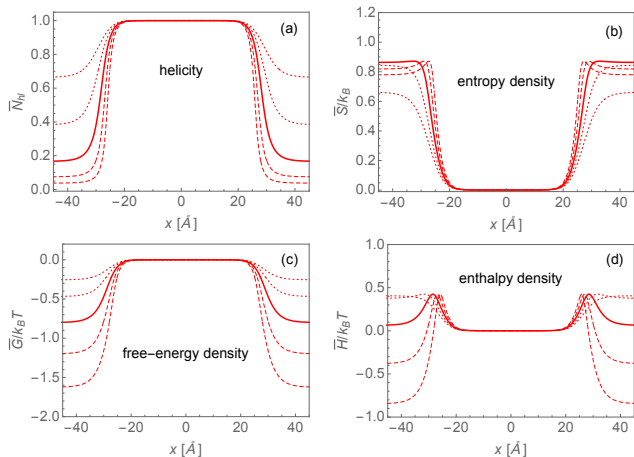
FIG. 5: Profiles of (a) helicity, (b) entropy density, (c) free-energy density, and (d) enthalpy density for a long polypeptide. The model parameter values are $\mu = 2$ and $\tau = 0.5$. The growth parameter field $t(x)$ uses (8)-(9) with $\rho_{\mathrm{w}}(x)$ from Fig. 4. The solid, dashed, and dotted curves pertain to $\alpha_{\mathrm{H}} = 1$, $\alpha_{\mathrm{H}} = 1.05, 1.1$, and $\alpha_{\mathrm{H}} = 0.95, 0.9$, respectively.
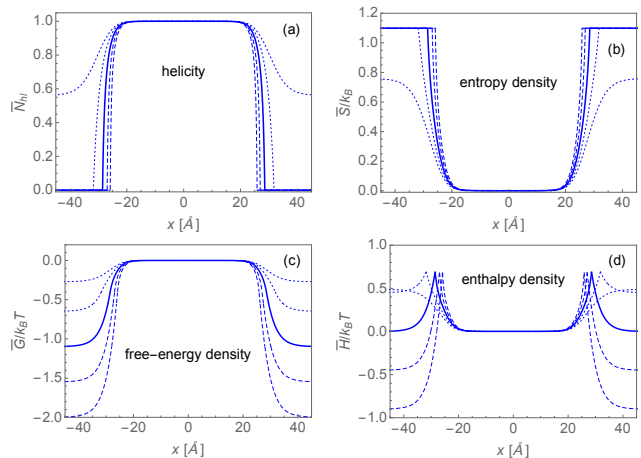
FIG. 6: Profiles of helicity, entropy density, free-energy density, and enthalpy density for a long polypeptide. The model parameter values are $\mu = \infty$ and $\tau = 0.5$. The growth parameter field $t(x)$ uses (8)-(9) with $\rho_{\mathrm{w}}(x)$ from Fig. 4. The solid, dashed, and dotted curves pertain to $\alpha_{\mathrm{H}} = 1$, $\alpha_{\mathrm{H}} = 1.05, 1.1$, and $\alpha_{\mathrm{H}} = 0.95, 0.9$, respectively.

In consequence, the free-energy density of the polypeptide rises only imperceptibly above its (zero) reference value as well. At positions away from the center of the bilayer, the helicity decreases and the entropy density increases, the former reflecting a drop in (helical) order and the latter a rise in (coil-like) disorder, both associated with the same conformational change.

These conclusions are not without caveat. Charged residues (e.g. Arg) tend to drag water into the membrane. Aggregates of peptides can form a polar environment inside the membrane in different ways, starting with water bridges and ending in pores, for example [80–82]

In Ref. [24] we identified one source of order and two sources of disorder involving the secondary structure of the polypeptide backbone alone. Order increases with the growth of segments of helix conformation. Disorder is contained (i) in the spatial distribution of boundaries between segments of coil and helix and (ii) inside each (unstructured) coil segment.

The enthalpy density produces an energy barrier at locations near the lipid-water interface [Figs. 5(d) and 6(d)]. Near the peak position the thermal fluctuations are just strong enough to break internal H-bonds, but the environment is not yet sufficiently polar to produce an adequate supply of external H-bonds as replacements. The enthalpy density decreases (for different reasons) on each side of the peak. On one side, broken bonds are rare, on the other side, they are energetically inexpensive. Outside the interface, the enthalpy profile levels off in a high or low plateau depending on the value of the physical parameter α_{H}. That parameter also affects the drop in helicity and the rise in entropy. Enthalpic loss ($\alpha_{\mathrm{H}} < 1$) favors order and suppresses disorder (for the peptide) in the aqueous environment.

There are some qualitiative and some quantitative differences between the curves for $\mu = 2$ and for $\mu = \infty$. The case $\mu = \infty$ produces pure coil with maximum entropy in the aqueous environment. Pure coil conformation means zero helicity. Coil segments generate significantly more entropy for $\mu = \infty$ than for $\mu = 2$. The enthalpic barriers near the interface are more pronounced in the case $\mu = \infty$. This difference is attributable to an entropic effect. The breaking of an H-bond at significant enthalpic cost is more likely to happen if the entropy produced is large ($\mu = \infty$) than if it is small ($\mu = 2$).

The free-energy profiles in Figs. 5 and 6 tell us that the incentives for the insertion of peptides must come from a source other than what has already been taken into account. The free-energy density $\bar{G}(x)$ is significantly higher in the membrane environment than in the surrounding water. The dependences of $\bar{H}(x)$ and $\bar{S}(x)$ on α_{H} are strong, but not decisively so. The entropic contributions accounted for thus far are dominating the free energy. Coil is favored over helix. Water beats lipids as the favored environment for the peptide.

B. External H-bonds

At this point in the analysis, we add one other contribution to the free-energy density profile. Further contributions, associated with side chains and their interaction with lipids, will be introduced in Sec. IV C. The contribution discussed here is entropic in nature and favors insertion. The replacement of backbone internal H-bonds with external H-bonds that immobilize H_2O molecules from the aqueous environment, while providing an enthalpic discount as already accounted for, comes at an entropic cost [83]. This effect is taken into account via an amended free-energy density constructed as follows:

$$\frac{\bar{G}_{\mathrm{H}}(x)}{k_{\mathrm{B}}T} = \frac{\bar{G}(x)}{k_{\mathrm{B}}T} + 2\frac{|\Delta\bar{S}_{\mathrm{H}}|}{k_{\mathrm{B}}}\Big[1 - \bar{N}_{\mathrm{hl}}(x)\Big], \qquad (11)$$

with $|\Delta S_{\mathrm{H}}|$ as discussed and estimated in Sec. II D. The factor in square brackets represents the fraction of backbone segments in coil conformation with each segment offering docks for two H_2O molecules. Profiles for the amended free-energy densities are shown in Fig. 7.

This amendment does indeed contribute an incentive for insertion. Its impact depends on the size of $|\Delta\bar{S}_{\mathrm{H}}|/k_{\mathrm{B}}$, which controls the entropy reduction due to external H-bonds, and on the range parameter μ, which controls entropy production of coil segments. An increase of $|\Delta\bar{S}_{\mathrm{H}}|/k_{\mathrm{B}}$ switches the global minimum in the free-energy density from the exterior region to the interior region. The switch happens at $|\Delta\bar{S}_{\mathrm{H}}|/k_{\mathrm{B}} \simeq 0.5$ for $\mu = 2$. A somewhat larger value, $|\Delta\bar{S}_{\mathrm{H}}|/k_{\mathrm{B}} \simeq 0.55$ is needed to cause the switch for $\mu = \infty$. These values are well below the estimated upper bound (5). Unsurprisingly, the larger range parameter, which produces more backbone entropy, requires more entropy reduction in compensation before it yields an insertion incentive.
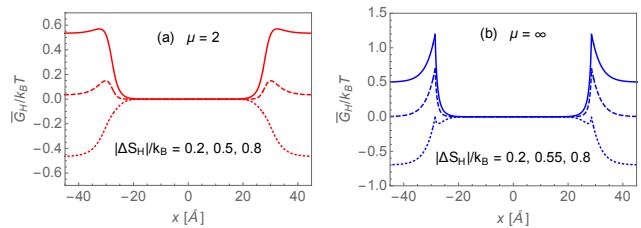

FIG. 7: Profiles of free-energy density (11) which includes an entropic contribution from external H-bonds. The two panels pertain to different values of the range parameter μ. The curves from bottom to top are for increasing values of $|\Delta\bar{S}_{\mathrm{H}}|/k_{\mathrm{B}}$. Additional model parameter values are $\tau = 0.5$ and $\alpha_{\mathrm{H}} = 1$. Note the different vertical scales left and right.

Interestingly, an energy barrier between the exterior and interior levels of free-energy-density builds up as $|\Delta\bar{S}_{\mathrm{H}}|/k_{\mathrm{B}}$ increases in size. This barrier, which is more pronounced in panel (b), helps stabilize the (experimentally confirmed) coexistence of short peptides in states of solution and adsorption. Energy barriers such as emerge here quite naturally, play an important role in kinetic studies of the insertion process.

IV. LANDSCAPES

In the continuation of the analysis, we interpret the profiles calculated in Sec. III as propensities of backbone segments of short peptides such as pHLIP. The helicity profile and the density profiles for enthalpy, entropy, and free energy are employed here as one factor affecting the behavior of residues of short peptides in the same environment, specifically their conformational preference (coil or helix). Other factors depend on attributes of the specific side chains and on further (enthalpic and entropic) effects of peptide-lipid interactions.

Landscapes of interest include those of free energy, enthalpy, entropy, and helicity. It would take a molecular dynamics simulation to explore entire landscapes, i.e. the full range of configurations that a peptide can assume in the membrane environment. The approach taken here, which is more limited in scope (see Sec. V), is more selective with configurations. This selectivity is justifiable, at least in part, by the fact that the range of configurations is naturally and severely restricted by conformational constraints, specifically by the rigidity of the helix conformation.

A. Three-variable landscapes

In the following, we investigate three-variable free-energy landscapes for short peptides in varying positions and orientations. The model peptide has N_{R} residues and is assumed to consist of two straight segments as

schematically represented in Fig. 8. In applications to pHLIP, a likely candidate for the kink position is the helix inhibiting Pro residue [45–47]. One simulation study [31] places a kink at the position of the Asp residue which is somewhat closer to the N terminus than the Pro residue.
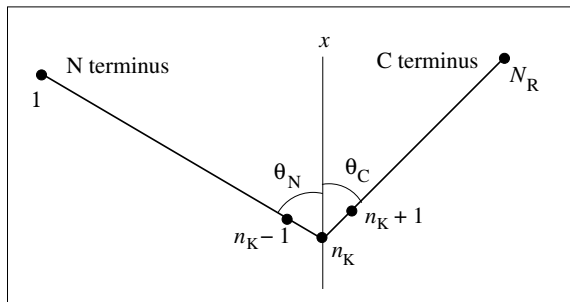

FIG. 8: Variables x_{n_K} (depth of kink) and θ_N, θ_C (angles between normal and segments). The counting of residues begins at the N terminus as is custom.

We use the position x_{n_K} of the kink on the normal to the bilayer and the angles θ_N, θ_C of the segments ending in the N, C terminus, respectively, as the variables that specify the position and orientation of the peptide in the membrane environment. The lengths of both segments depend on the local conformation, which, in turn, depend on the location of a given segment in the membrane environment as specified below.

We intend to calculate the free energy of peptides along specific pathways in this three-variable landscape. Our focus will be on the pHLIP variants with the sequences as stated in Fig. 1. We consider three principal states I, II, and III of pHLIP in the membrane environment and specific pathways between them [9, 26]:

I: pHLIP is in aqueous solution and in coil conformation. We shall start adsorption pathways at coordinates $x_{n_K} = 27$Å, $\theta_N = \theta_C = 90°$.

II: pHLIP is adsorbed to the outside interface of a cell membrane with interstitial fluid or of a liposome with water. At high pH the adsorption is rather loose and the conformation is coil. At low pH the adsorption is deeper and the conformation is largely α-helix. We shall see that the adsorbed state has the kink of the peptide positioned deeper in the membrane and both termini sticking out into the water: $x_{n_K} \simeq 17$Å, $\theta_N \simeq \theta_C \simeq 30°$.

III: pHLIP is in a trans-membrane state with a helical central part and short coil-like flanking parts. The coordinates roughly are $x_{n_K} \simeq 0$, $\theta_N \simeq 0$, $\theta_C \simeq 180°$.

In order to calculate the peptide free energy in any of the three principal states and in states that connect them we need to know the position x_n of every residue in the membrane environment. We determine these positions recursively from distances between residues beginning at the kink. Writing,

$$x_{n-1} = x_n + l(x_n)\cos\theta_N \quad : n = n_K,\ldots,2, \qquad (12)$$
$$x_{n+1} = x_n + l(x_n)\cos\theta_C \quad : n = n_K,\ldots,N_R - 1,$$

takes into account that the distance between adjacent residues depends on the local conformation of the residue position in the membrane environment. That conformation is either coil or helix with probabilities for which we use the propensity profile as calculated in Sec. III:

$$x_{n+1} - x_n \doteq l(x_n) = l_c - (l_c - l_h)\bar{N}_{hl}(x_n). \qquad (13)$$

Whereas the length of a helical segment is given, $l_h = 1.5$Å, the (averaged) length l_c of coil segments becomes a physical model parameter. Its maximum value is the contour length per residue of the backbone: $l_c \lesssim 4$Å.

B. Backbone contributions

It is useful to look at some free-energy and helicity landscapes that represent the effects of the backbone alone. This is facilitated in Appendix B. The main message from the results presented there to what follows is twofold: (i) The backbone-lipid interaction favors coil conformation near the water-lipid interface and helix conformation inside the membrane. (ii) Insertion lowers the backbone contribution to the free energy.

Here we merely state the expression for the backbone contribution to the free-energy landscape and the expression for the helicity (fraction of peptide in helix conformation) to be used in the following:

$$G_{BB}(x_{n_K}, \theta_N, \theta_C) = \sum_{n=1}^{N_R} \bar{G}_H(x_n), \qquad (14)$$

$$N_{hl}^{(K)}(x_{n_K}, \theta_N, \theta_C) = \sum_{n=1}^{N_R} \bar{N}_{hl}(x_n), \qquad (15)$$

where the function $\bar{G}_H(x)$ is taken from (11) and the functions $\bar{G}(x)$, $\bar{N}_{hl}(x)$ from Appendix A. The dependence of x_n on $x_{n_K}, \theta_N, \theta_C$ is given in (12).

C. Side-chain contributions

The contributions to the free-energy landscape originating from the side chains of a peptide with a given sequence of residues are manifold [84]. Their interactions with lipid molecules include aspects of hydrophobicity, pressure differentials, entropy reductions, and electrostatics. Ranking the relative importance has been challenging and not without controversy. For the adsorption and insertion pathways presented in Sec. V we use well-established transfer-free-energy data as the dominant side-chain contribution. The potential impact of other contributions will be discussed in Sec VI as an outline of future work.

The side chains of pHLIP residues range from strongly hydrophobic to strongly hydrophilic. Their transfer between polar and nonpolar environments contributes significantly to the free-energy landscape. Some side chains

carry (positive or negative) electric charges. The protonation at low pH of the negatively charged Asp residues and C terminus (see Fig. 1) changes the overall hydrophobicity critically as we shall see, destabilizing the adsorbed state II in favor of the trans-membrane state III. An estimate of the side-chain transfer-free energy between states I, II, and III based on the Wimley-White interface and octanol scales is given in Appendix C. That scheme takes into account free-energy differences between three levels, w, i, o, representing (polar) water, (mixed) interface, and (nonpolar) octanol or lipid-hydrocarbon-tail environments, respectively. Our modeling with these data strongly indicates that the aforementioned instability is real and in agreement with experiments [29, 32, 41].

For the investigation of insertion pathways we replace the three environmental levels w, i, o by an environmental field in the manner discussed in Sec. II. We use the density field of water (1) to specifically convert the steps $\Delta G_{\mathrm{wo}}^{(n)}$ from Wimley-White data [85–91] for the residue at sequential position n into fields of gradual change:

$$G_{\mathrm{res}}^{(n)}(x) = \Delta G_{\mathrm{wo}}^{(n)}\left[1 - \rho_{\mathrm{w}}(x)\right]. \qquad (16)$$

The side-chain contribution to the free-energy of the peptide then becomes,

$$G_{\mathrm{SC}}(x_{n_{\mathrm{K}}}, \theta_{\mathrm{N}}, \theta_{\mathrm{C}}) = \sum_{n=1}^{N_{\mathrm{R}}} G_{\mathrm{res}}^{(n)}(x_n), \qquad (17)$$

with the dependence of x_n on $x_{n_{\mathrm{K}}}, \theta_{\mathrm{N}}, \theta_{\mathrm{C}}$ from (12). From here on we use the specifications $x_{\mathrm{a}} = 17.8$Å, $x_{\mathrm{b}} = 26.3$Å, $x_{\mathrm{s}} = 2.0$Å for the profile (1).

The transfer free energies relative to an exterior position for the hydrophobic Val residue and the mildly hydrophilic Asp residue (when protonated) are shown in Fig. 9 (solid curves) as an illustration of this scheme. The deprotonated Asp residue is much more strongly hydrophilic, which costs significant extra free energy, as indicated by the dotted line. The dashed line represents a modification of the Asp profile that takes into account the complication that during the insertion process, the pH is different in the exterior and interior regions. This last point requires some explanation.

When the pH is lowered in the exterior region, all Asp residues quickly become protonated and, in consequence, less hydrophilic. Their transfer-free-energy profile changes from the dotted line to the solid line in Fig. 9. It is well established in experiments with liposomes that the insertion process is faster than the equilibration of the pH across the membrane. Hence the three Asp residues closest to the C terminus will only remain protonated, during the insertion process, until they come into contact with the aqueous environment in the interior region, which is still at higher pH. The consequence in the framework of quasistatic translocations considered here is that the transfer-free energies of these Asp residues drop to a level below that of the exterior region, as is reflected in the dashed profile of Fig. 9. As the pH slowly drops inside the liposome, the modified (dashed) profile gradually
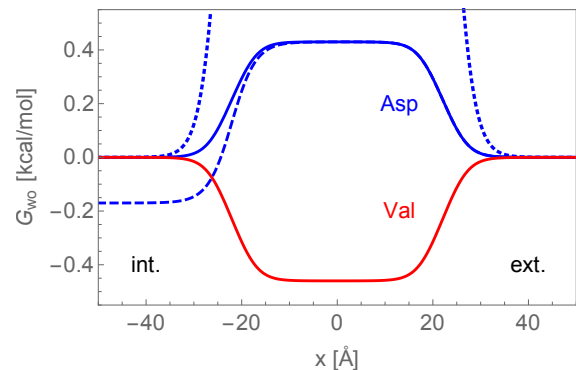

FIG. 9: Transfer-free-energy profile of one hydrophobic residue, Val, and one hydrophilic residue Asp, when protonated (solid line) or deprotonated (dashed line). The dashed line represents a hybrid profile for the same Asp residue, which accounts for a kinetic effect as described in the text.

turns back into the profile characterizing the protonated Asp (solid line).

In the projected kinetic study of peptide insertion [39], this complication will, of course, be accounted for in a very different way. The goal here is more modest. We aim to identify the circumstances under which insertion pathways for pHLIP in the free-energy landscape exist. There are no intrinsic time scales associated with these pathways. However, it appears to be a necessary condition for insertion, as will be explained in Sec. V, that a differential in pH between exterior and interior regions is maintained. In liposome experiments, this pH differential only exists for a limited time, during which insertion must be completed.

V. PATHWAYS

Here we pick up the thread from Sec. IV A, where we designed a template for exploring adsorption and insertion pathways of a peptide consisting of two straight segments with a kink at the Pro position (our choice). Any pathway taken by the peptide must be downhill in the free-energy landscape, assembled from the backbone part (14) and the side-chain part (17),

$$G(x_{n_{\mathrm{K}}}, \theta_{\mathrm{N}}, \theta_{\mathrm{C}}) = G_{\mathrm{BB}}(x_{n_{\mathrm{K}}}, \theta_{\mathrm{N}}, \theta_{\mathrm{C}}) + G_{\mathrm{SC}}(x_{n_{\mathrm{K}}}, \theta_{\mathrm{N}}, \theta_{\mathrm{C}})$$
$$= \sum_{n=1}^{N_{\mathrm{R}}} \left[\bar{G}_{\mathrm{H}}(x_n) + G_{\mathrm{res}}^{(n)}(x_n)\right], \qquad (18)$$

with x_n from (12).

We have adopted a very simple search procedure for descending pathways. For each step we consider positive and negative infinitesimal increments for the three variable $x_{n_{\mathrm{K}}}, \theta_{\mathrm{N}}, \theta_{\mathrm{C}}$ (see Fig. 8) and execute the step that produces the most negative ΔG. The pathway stops when none of the variable changes generates a $\Delta G < 0$.

Guided by the circumstances associated with experimental investigations of pHLIP insertion into liposomes, we divide the pathways toward lower free energy into three consecutive phases. Different environmental conditions pertain to each phase.

(A) During the *adsorption* phase, the pH is high and all Asp residues are deprotonated. The peptide starts out in solution close to the membrane. Pathways with $\Delta G < 0$ end with the peptide adsorbed at the interface of the membrane with the exterior aqueous region.

(B) The *insertion* phase starts with the peptide in that adsorbed state but with the environment changed. The pH is low in the exterior region and all Asp residues have become protonated. Pathways with $\Delta G < 0$ end with the peptide in a trans-membrane state. The pH remains high in the interior region.

(C) The *stabilization* phase comes into play once the pH has dropped in the interior region (of a liposome), which typically happens more slowly than the insertion. It affects the stability of the trans-membrane state. However, pathways with $\Delta G < 0$ keep the peptide inserted.

A. Adsorption

The gray background of Fig. 10 is a representation of the water-density profile [Fig. 4(a)] across the heterogeneous membrane environment. The exterior region is at the top. At the starting point of the pathway explored here, the adsorption phase is already on its way. The peptide is sprawled wide near the interface.
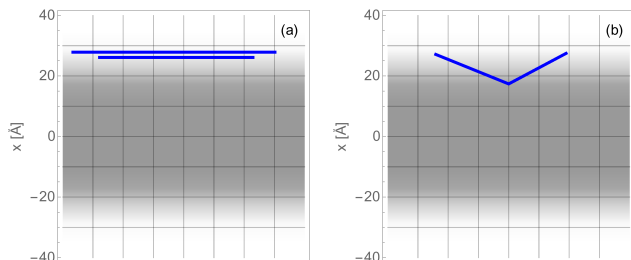

FIG. 10: Variant W6 of pHLIP modeled as two straight segments with a kink at the Pro position during the adsorption phase of the pathway.

The adsorption pathway from here on consists of two legs. First the peptide moves toward the membrane without changing its orientation. Along this stretch, it undergoes a conformational change from coil to mostly helix, which shortens its length. The initial and final configurations of this first leg are shown in Fig. 10(a). Along the second leg of the adsorption pathway, the two segments on either side of the Pro kink change their orientations and the kink position continues to move toward the center of the membrane. This brings the hydrophobic residues deeper into the nonpolar environment, yet keeps the hydrophilic residues closer to the polar environment. The final, adsorbed equilibrium state is shown in Fig. 10(b).

Additional information of a more quantitative nature about the adsorption pathway is compiled in Fig. 11. The two legs are clearly discernible in the variations of the kink position [panel (a)] and the angles of orientation of the two segments [panel (b)]. The first leg ends after step $i = 8$. Here the angles begin to vary and the kink moves at a slower rate, meaning down a more shallow slope. The pathway ends after step $i = 100$. The two legs are more clearly recognizable in some data than in others. The zigzag lines visualize the search protocol described earlier.
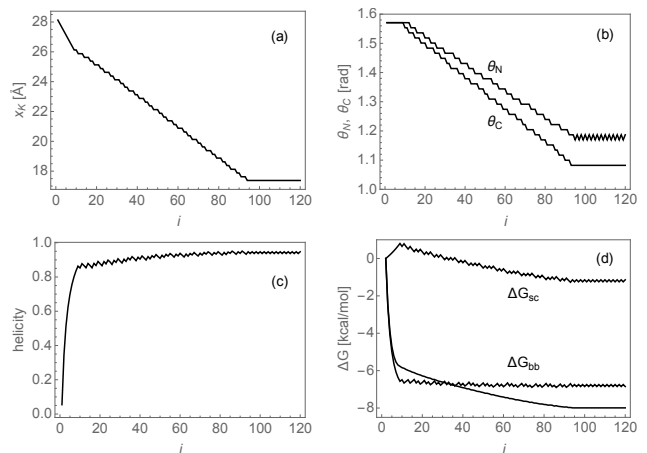

FIG. 11: (a) Kink position, (b) angles (away from vertical orientation) of the two segments, (c) helicity, (d) change in free energy (backbone, side-chain contributions and total) versus step number i along the adsorption pathway.

The conformational change from coil to helix happens almost entirely during the first leg. Here the decrease in free energy is driven by the backbone contribution. The conformational change, which involves the backbone, takes place during the first leg. An environmental differentiation restricted by conformational constraints takes place during the second leg. Hydrophobic residues (near the center of the peptide) move further into membrane while hydrophobic residues stay closer to water. We have chosen the starting point of the adsorption phase such that the overall change of free energy (~ -6 kcal/mol) is consistent with caloric experiments [2, 28].

B. Insertion

When the pH is lowered in the exterior region, the final state of the adsorption pathway becomes unstable. It becomes the initial state of the next phase of the pathway [Fig. 12(a)]. All Asp residues are now protonated [27, 92]. The destabilizing agent is the switch of their transfer-free-energy profile from the deprotonated to the
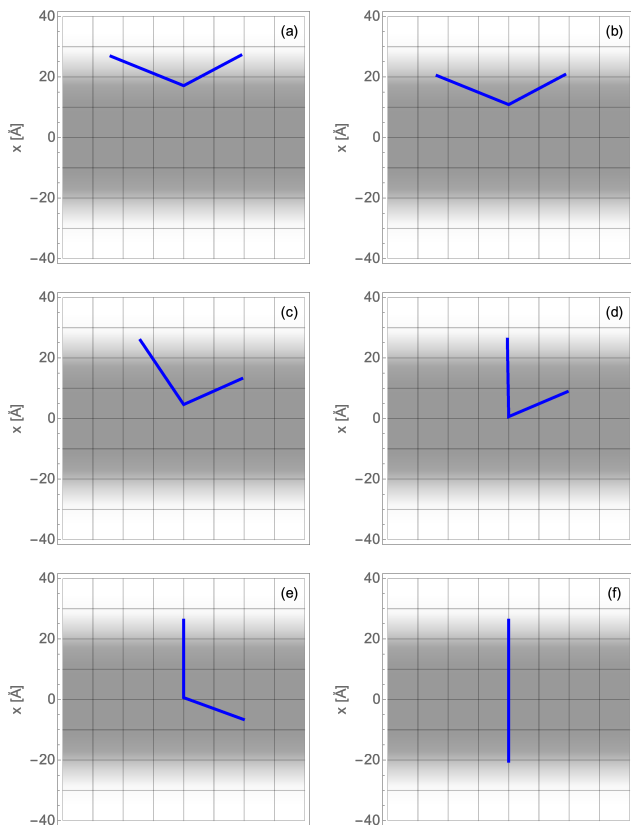
FIG. 12: Variant W6 of pHLIP modeled as two straight segments with a kink at the Pro18 position during the insertion phase of the pathway.

hybrid version. This modification opens up pathways of descending free energy, which we explore using the previously described protocol. What emerges turns out to be an insertion pathway. It consists of three legs, as visualized in Fig. 12.

Along the first leg [panels (a)-(b)], the peptide sinks somewhat deeper into the membrane with no significant change in the orientation of the two segments on either side of the kink at the Pro position. The protonation has made the Asp residues significantly less hydrophilic. This has weakened their resistance against the pull of the hydrophobic residues into the nonpolar environment. Hence the downward translocation of the entire peptide.

Along the second leg [panels (b)-(d)], the segment on the left with the N terminus at its end straightens up while the other segment does not rotate significantly. The Pro kink position moves further toward the center of the membrane. Protonation has taken place primarily near the C terminus. The positive charges at the N terminus and at the Arg residue on the same segment are still present. Reorienting that segment into a transmembrane direction keeps those charges in or near the polar environment and allows the hydrophobic residues to move deeper into nonpolar environment.

The third leg of the insertion pathway [panels (d)-(f)] is mainly driven by the forces acting on the segment with the C terminus at its end. There are competing forces in action. The hydrophobic forces lead the initial descent in free energy toward a horizontal orientation of that segment. The subsequent descent in free energy toward the trans-membrane orientation is guided by the forces acting on the protonatable Asp residues and the C terminus. Recall that the interior region is still at high pH, which significantly enhances the pull of the protonatable contacts toward the nonpolar environment, where deprotonation takes place.

Figure 13 provides additional, more quantitative information about the insertion pathway. The three legs are most clearly discernible in the kink position and the two angles of orientation. The reorientation of the two segments during the second and third legs is almost completely sequential: first the N terminus backs out into the exterior region while the kink position moves toward the center; then the C terminus moves across into the interior region while the kink position remains stationary.

The helicity (already strong) reaches near saturation during the first leg, but then decreases somewhat as the N terminus reestablishes closer contact with the exterior aqueous region. The change in free energy during insertion originates predominantly from the side chains. The drop of backbone free energy is less than 0.3 kcal/mol and happens quickly while the helicity increases. The drop in side-chain free energy is, for the most part, spread across the first two legs of the insertion pathway. Its amount of ~ 4kcal/mol is a bit higher than what caloric experiments predict, but not by much [2, 28].

In our study we have worked with a single density field of water, not taking into account any possible variations due to a change in pH. We did confirm though that the insertion pathways described above are robust under small
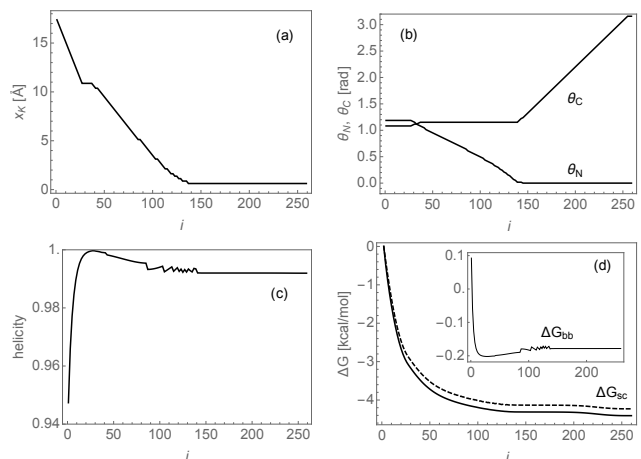

FIG. 13: (a) Kink position, (b) angles (away from vertical orientation) of the two segments, (c) helicity, (d) change in free energy (backbone, side-chain contributions and total) along the insertion pathway.

changes of the parameters x_a, x_b. Variations up to 5% in either parameter did not produce any qualitative changes in the insertion pathway and the stability of the transmembrane state. The only proviso is that for smaller x_a, the initial position of the peptide must be closer to the membrane for adsorption to ensue.

C. Stability

In an experiment that uses liposomes instead of biological cells, the low level of pH imposed on the exterior region will slowly leak into the interior region. This happens on a significantly slower time scale than the insertion process. There is clear experimental evidence that pHLIP stays inserted as the pH equilibrates at a low level. Hence the third phase of the pathway must ensure that when we change the Asp transfer-free-energy profiles from the hybrid version to the protonated version the transmembrane state remains stable [26, 30].

This turns out to be the case indeed as illustrated in Fig. 14. Panel (a) shows the final state of the second phase and panel (b) the final state of the third phase. The low pH in the interior region reprotonates the C terminus and the Asp residues near it. Hence their counteraction against the hydrophobic forces on the Leu residues (also near the C terminus) weakens. This has the effect that the segment with the C terminus sticking into the interior region changes its orientation somewhat to find the new local free-energy minimum.
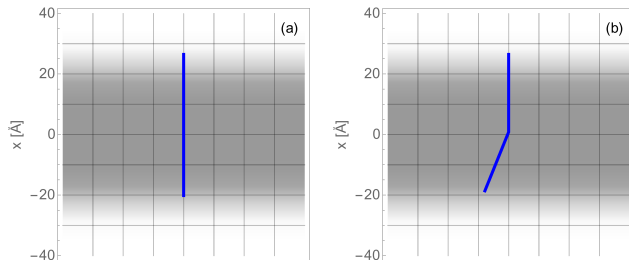

FIG. 14: Variant W6 of pHLIP modeled as two straight segments with a kink at the Pro18 position during the third phase of the pathway.

The kink position and the orientation of the segment with the N terminus at its end do not undergo significant changes. The descent in free energy is small in comparison to the previous two stages. It amounts to less than 10cal/mol. The important message is that the peptide stays in a trans-membrane state. Note that insertion (second stage) only happens, according to the protocol of our pathway exploration, if the pH remains high in the interior region until the trans-membrane state has been realized. However, once the trans-membrane state is realized it remains stable even after the pH has dropped in the interior region.

Let us emphasize that Figs. 11 and 13 are not meant to suggest any times scales for the adsorption and insertion processes, respectively. Our protocol for the exploration of free-energy landscapes merely searches for downhill directions and then takes steps accordingly with no attempts undertaken to optimize directions for steepest descent. A kinetic study of the pHLIP adsorption and insertion processes is an entirely different project [39].

VI. CONCLUSION AND OUTLOOK

Processes of peptide insertion into a membrane are highly complex in their kinetics and can be very diverse in their energetics. In this work an effort has been described that focuses on the energetics of a particular scenario. The specific goal has been to demonstrate plausible insertion pathways in a free-energy landscape assembled from contributions identified as dominant. We built that landscape from enthalpic and entropic contributions associated with internal and external H-bonds along the peptide backbone and from transfer free energies of side chains between hydrophilic and hydrophobic regions of the membrane environment.

The advantage of the approach centered on energetics is that the problem associated with the inevitably broad range of time scales has come into play only marginally. The investigation of free-energy landscapes and the search for insertion pathways is not meant to be a substitute for a molecular dynamics simulation or other form of kinetic study [31, 55, 68, 93–101]. It is an exploration of the circumstances under which kinetic processes associated with peptide insertion are likely to take place.

Insertion pathways as explored in this work are characteristic of quasistatic processes. The free-energy landscapes provide road signs, but no time tables. The road signs thus established are not infallible because the free-energy landscapes in use are manifestly incomplete. A strong case can be made that what has been left out is more productively taken into account in a kinetic study or a molecular dynamics simulation. The most relevant missing pieces represent different aspects of lipid-peptide interactions. Some are predominantly enthalpic and other predominantly entropic in nature.

(i) The lateral pressure profile of the membrane, which provides mechanical stability to the bilayer structure against perturbations of various kinds, is known to have a characteristic shape with regions of positive deviations from ambient pressure sandwiching a narrow band of negative deviation just inside the lipid headgroups. These empirically established pressure variations are significant [60–69].

(ii) In the (trans-membrane) state III the predominently hydrophobic and α-helical center of pHLIP is flanked by coil segments that are more hydrophilic. In a positive (negative) mismatch, the hydrophobic center is too long (short) to be comfortably accommodated in the membrane. Each type of mismatch elicits a distinct

response from both the peptide and the membrane. In a positive mismatch, the membrane tends to increase its width locally. Sheltering additional hydrophobic residues from water costs elastic energy. That cost can be lowered if the helical axis tilts its trans-membrane orientation away from the bilayer normal. To a negative mismatch the membrane responds with a local thinning, which puts the helix under tension. Some of that tension may be released by a partial change from α-helix to the more tightly wound 3_{10}-helix [102–108].

(iii) The lipid molecules in the bilayer, typically bent into a liposome of spherical shape, are in a macrostate which represents a two-dimensional fluid. The positional ordering of the headgroups is then of short range over some coherence length ξ. The undisturbed lipid bilayer is characterized by a uniform entropy density with two contributions, one representing positional disorder of the headgroups and the other orientational and conformational disorder of the hydrocarbon tails [11, 109]. Headgroups and tails perform a free-energy balancing act of sorts. An increase in bilayer width produces more space for the tails to explore and requires work by the headgroups against the pressure of the surrounding water, thus implying $T\Delta S > 0$ and $\Delta H > 0$, respectively. A decrease in the average number of headgroups per unit area increases their positional freedom, implying $T\Delta S > 0$, but compresses the tails, implying $\Delta H > 0$. Balance is reached when $\Delta G = \Delta H - T\Delta S = 0$.

(iv) The presence of pHLIP produces a contact line with lipid molecules. Along that contact line, the aforementioned thermal equilibrium of headgroups and tails is being disturbed [110, 111]. The dominant changes involve an entropy reduction in both headgroups and tails. The former tend to line up tightly against the foreign object and form an ordered layer of width roughly equal to the aforementioned coherence length ξ. pHLIP going into the (adsorbed) state II throws the normal pressure out of balance. Fewer tails must exert the same force per unit area. They can do that only under higher compression, facilitated by membrane thinning as noted before. This, in turn, leads to an entropy reduction. When pHLIP inserts the contact line becomes much shorter. The associated decrease in ΔG dominated by a positive $T\Delta S$ is a factor favoring insertion thermodynamically. Jähnig [10] investigated this effect under the name *lipohobic effect* and used a coherence length of $\xi \simeq 15$Å.

(v) What we have been calling water-lipid interface involves, in fact, an electrolyte with variable ion concentration on one side. In our study we have taken into account one particular aspect of this presence and variation, namely the effect of the pH on the protonation status of negatively charged residues. There are different, well documented ways in which ion content of the water affects the lipid bilayers, some of which overlap with the previous items of this list [112–119].

Taking into account effects (i)-(v) calls for a more detailed model of the peptide than is being used in this study. The stochastic modeling intended to be used for pHLIP insertion kinetics [39] is a Markov chain model, which promises interesting points of comparison with molecular dynamics simulations. One strong point of Markov chain modeling is that it can deal with a range of different time scales in a most transparent and efficient way. The price to be paid for that advantage is that interactions on a microscopic scale are accounted for more summarily than a simulation does at the molecular dynamics level.

Appendix A: Statistically interacting polymer links

The origin in quantum many-body theory [120–124] of the methodology used here and its adaptation to problems of current interest in classical statistical mechanics [37, 125–130] has already been presented from several different angles. Among these works are papers dedicated to polypeptides [24] and to double-stranded DNA [131].

Here we summarize those results from Ref. [24] which are being used (first in Sec. III) as the main building blocks for extensions reported in this work. The microscopic model for the coil-helix transition of a long polypeptide at a water-lipid interface solved in [24] has three parameters: the growth parameter t, the nucleation parameter τ, both continuous, and the (discrete) range parameter μ, which numbers the coil states available to each residue (see Fig. 3). Here we consider the cases $\mu = 2$ and $\mu = \infty$ at $\tau > 0$ and use selected results. We quote the relevant expressions in ways easy to trace back to their derivations.

The scaled Gibbs free energy in closed form reads

$$\frac{\bar{G}(t,\tau)}{k_{\mathrm{B}}T} = -\ln\left(\frac{1+w(t,\tau)}{t}\right), \qquad (A1)$$

where, for $\mu = 2$,

$$w(t,\tau) = \frac{1}{3}\left[x + 2\sqrt{x^2 + 3y}\cos\frac{\varphi}{3}\right], \qquad (A2)$$

$$\tan\varphi = \frac{\sqrt{27(4y^3 + y^2x^2 + 18yx^2 + 4x^4 - 27x^2)}}{x(2x^2 + 9y - 27)},$$

with $x \doteq t-1$, $y \doteq 1+t\tau$ and $0 \leq \varphi < \pi$, and, for $\mu = \infty$,

$$w(t,\tau) = \begin{cases} 2 & : 0 \leq t \leq t_{\mathrm{c}}, \\ t-1 + \dfrac{t\tau}{\lambda} & : t > t_{\mathrm{c}}, \end{cases} \qquad (A3)$$

with

$$\lambda(t,\tau) \doteq \frac{1}{2}\left[t-1 + \sqrt{(t+1)(t-3) + 4t\tau}\right] \qquad (A4)$$

and

$$t_c \doteq \frac{3}{1+\tau}. \qquad (A5)$$

The expressions for helicity (order parameter) and entropy, inferred from first derivatives of $\bar{G}$ become

$$\bar{N}_{\mathrm{hl}}(t,\tau) = \frac{t}{1+w}\frac{\partial w}{\partial t}, \qquad (A6)$$

$$\frac{\bar{S}(t,\tau)}{k_{\mathrm{B}}} = \ln(1+w) - \frac{1}{1+w}\left[t\ln t\frac{\partial w}{\partial t} + \tau\ln\tau\frac{\partial w}{\partial\tau}\right], \qquad (A7)$$

respectively. They are too unwieldy for $\mu = 2$ to be reproduced here but fairly concise for $\mu = \infty$:

$$N_{\mathrm{hl}}(t,\tau) = \begin{cases} 0 & : 0 \le t < t_c, \\ 1 - \dfrac{t\tau}{\lambda^2 - 1 + t\tau} & : t \ge t_c, \end{cases} \qquad (A8)$$

$$\frac{\bar{S}(t,\tau)}{k_{\mathrm{B}}} = \ln\left(1 + \frac{\tau}{\lambda}\right) + \frac{t\tau}{\lambda^2 - 1 + t\tau}\left(\ln t - \frac{\lambda^2 - 1}{t(\lambda + \tau)}\ln\tau\right). \qquad (A9)$$

Finally the scaled enthalpy,

$$\frac{\bar{H}(t,\tau)}{k_{\mathrm{B}}T} = \frac{\bar{G}(t,\tau)}{k_{\mathrm{B}}T} + \frac{\bar{S}(t,\tau)}{k_{\mathrm{B}}}$$
$$= \frac{t\tau}{\lambda^2 - 1 + t\tau}\left(\ln t - \frac{\lambda^2 - 1}{t(\lambda + \tau)}\ln\tau\right), \qquad (A10)$$

is of importance in this work.

Appendix B: Backbone effects

It is instructive to take a look at the backbone contributions to the landscapes of helicity and free energy for the case of fixed angles $\theta_{\mathrm{N}} = 0$, $\theta_{\mathrm{C}} = \pi$, notwithstanding its limitations. The simplest case replaces the second Eq. (13) by

$$l(x) = l_{\mathrm{h}} = 1.5\text{Å}, \qquad (B1)$$

implying that the position coordinate of successive residues progresses uniformly and independently of conformation. The value chosen in (B1) is accurate for a helical segment but shorter than most coil segments.

Free-energy and helicity landscapes thus predicted are shown in Fig. 15 for $\mu = 2, \infty$ and $N_{\mathrm{R}} = 35, 47, 23$. The first row of landscapes ($N_{\mathrm{R}} = 35$) is tailored to represent some variant of pHLIP, the second row a significantly longer peptide ($N_{\mathrm{R}} = 47$), and the third row a significantly shorter peptide ($N_{\mathrm{R}} = 23$). We only consider enthalpy parameters that are neutral $\alpha_{\mathrm{H}} = 1$ or represent a gain $\alpha_{\mathrm{H}} = 1.05, 1.1$.
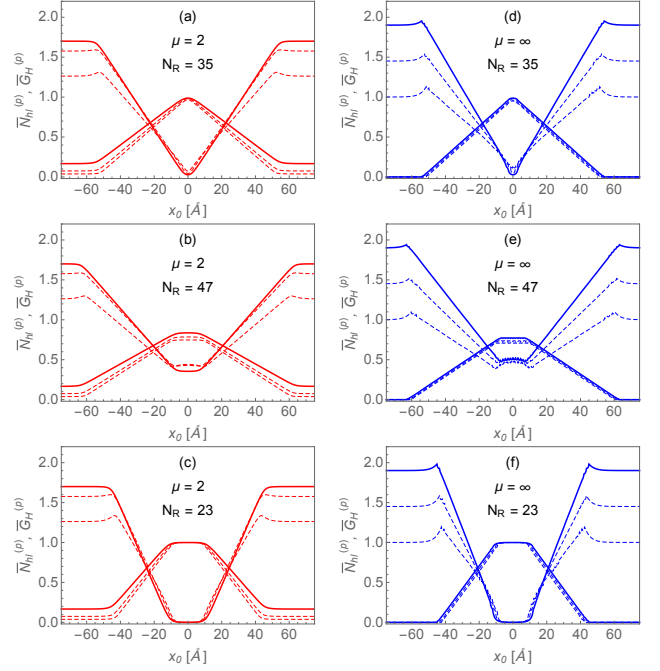

FIG. 15: Helicity (peaked at center) and scaled free-energy of peptide with N_{R} residues oriented and positioned as described in the text versus the coordinate x_0 of the central residue for $\mu = 2$ (left) and $\mu = \infty$ (right) and three sizes. The solid and dashed curves pertain to the values $\alpha_{\mathrm{H}} = 1$ and $\alpha_{\mathrm{H}} = 1.05, 1.1$ of the enthalpy parameter, respectively.

The case $\mu = \infty$ assigns more entropy to coil segments than the case $\mu = 2$. The effect on the results is significant, but produces no qualitative changes. The helicity landscapes are almost independent of the enthalpy parameter. That parameter affects the free-energy primarily in the aqueous environment as expected.

Insertion into the membrane is clearly favored in all three cases and for both variants of the model. The plots also tell us that insertion is accompanied by a conformational change from coil to helix. For the longest peptide the minimum in free energy is not as deep and the maximum in helicity is not as high as is the case for the two shorter ones. The obvious reason is that the former has significant flanking ends that remain in water.

Of particular interest is the free energy barrier that separates states with the center of the peptide in aqueous or membrane environments, the former mostly in coil conformation and the latter in helix conformation. This free-energy barrier is very shallow for $\mu = 2$ and only exists if $\alpha_{\mathrm{H}} > 1$. For $\mu = \infty$, on the other hand, it is more conspicuous and present even for $\alpha = 1$. This difference is related to the higher entropy that coil segments must shed if $\mu = \infty$ when they order into helix segments before they can cash in the enthalpic benefit of the lipid environment.

One message we take from this simplest case is that insertion is not automatic. An environmental change may

be needed to push the peptide over the barrier. An increase in acidity is known to do the trick. It neutralizes negatively charged residues (e.g. Asp and Glu) via protonation. The consequences for pHLIP are well documented by experiments [2, 26–28]. Let us recall that the extent of insertion as predicted by free-energy landscapes and the extent of ordering as predicted by helicity landscapes can be directly monitored experimentally, namely by Trp fluorescence and by circular dichroism experiments, respectively [30, 30, 31, 42].

An improved level of modeling takes into account that the distance between adjacent residues is different in the coil and helix conformations as reflected in Eqs. (13). The distance between successive residues now depends on the local conformation of the backbone at their position in the membrane environment. The main features in the results of the improved model (not shown) remain qualitatively the same.

Appendix C: Environmental levels

In a first round of estimates we may consider each residue placed in one of three distinct environments: polar (w), interface (i), or nonpolar (o). For the transfer free energies between any two environments we use the Wimley-White interface and octanol scales from [85–88]. State I (see Sec. IV A) has all residues in solution. For state II we assume that all residues are at the interface. With that assumption we use ΔG_{wi} for all residues to calculate the transfer free energy of the entire peptide. In this way we get one peptide transfer free energy,

$$\Delta G_{\mathrm{I-II}}^{\mathrm{hpH}} = +0.18\mathrm{kcal/mol}, \qquad (\mathrm{C1})$$

if the water is at high pH and another peptide transfer free energy,

$$\Delta G_{\mathrm{I-II}}^{\mathrm{lpH}} = -7.62\mathrm{kcal/mol}, \qquad (\mathrm{C2})$$

if the water is at low pH. In the former we assume that the negatively charged residues are deprotonated and in the latter we assume that they are protonated.

The more important result is (C1). It is well-established that at high pH pHLIP coexists in states I and II. Our rough estimate, which favors state I but only slightly, by a small fraction of $k_{\mathrm{B}}T$, is consistent with empirical evidence. A refined model will take into account that adsorbed pHLIP is largely in coil conformation. The mechanical flexibility of this conformation allows the hydrophobic residues to be closer to the interface and some hydrophilic ones to stick out into water. This will lower the free-energy prediction for state II relative to state I sufficiently to make $\Delta G_{\mathrm{I-II}}^{\mathrm{hpH}}$ negative but not nearly as much as (C2). A slightly negative value of $\Delta G_{\mathrm{I-II}}^{\mathrm{hpH}}$ is most consistent with experimental evidence.

Next we investigate, again by rough estimate, the peptide transfer free energies $\Delta G_{\mathrm{II-III}}$ associated with insertion. However, during the insertion process, from state II to state III, not all residues switch environment. We assume that in state III the trans-membrane part of the peptide is in α-helix conformation. The relevant bilayer width, ~ 35Å, then accommodates 23 residues in an α-helix conformation. These residues experience the transfer io. The five residues closest to the N terminus and the four residues closest to the C terminus are assumed to remain in the interface region. We thus obtain the following results for the variants W6 and W17 (see Fig. 1):

$$\Delta G_{\mathrm{II-III}}^{\mathrm{hpH}} = -0.5\mathrm{kcal/mol} \quad : \mathrm{W6}, \qquad (\mathrm{C3})$$

$$\Delta G_{\mathrm{II-III}}^{\mathrm{lpH}} = -4.32\mathrm{kcal/mol} \quad : \mathrm{W17}, \qquad (\mathrm{C4})$$

at high and low pH, respectively. The numbers are slightly lower for variant W30. The more important result is (C4). At low pH, the experimental evidence shows that insertion is clearly favored over the adsorbed state. We also see that the transfer free energy of insertion due to hydrophobic forces associated with side chains is considerably higher than the corresponding backbone transfer free energy investigated in Sec. IV. The result (C3) is significant for the pHLIP exit process (to be discussed elsewhere).

[1] O. A. Andreev, D. M. Engelman, and Y. K. Reshetnyak, pH-sensitive membrane peptides (pHLIPs) as a novel clss of deivery agents, Mol. Membr. Biol. **27**, 341 (2010).

[2] D. Weerakkody, A. Moshnikova, M. S. Thakur, V. Moshnikova, and J. Daniels, Family of pH (Low) Insertion Peptides for tumor targeting, PNAS, **110**, 5834 (2013).

[3] L. Yao, J. Daniels, D. Wijesinghe, O. A. Andreev, and Y. K. Reshetnyak, pHLIP-mediated delivery of PEGylated liposomes to cancer cells, J. Contr. Release **167**, 228 (2013).

[4] E. A. Sosunov, E. P. Anyukhovsky, A. A. Sosunov, A. Moshnikova, and D. Wijesinghe, pH (Low) Insertion Peptide (pHLIP) targets ischemic myocardium, PNAS

110, 82 (2013).

[5] O. A. Andreev, D. M. Engelman, and Y. K. Reshetnyak, Targeting diseased tissues by pHLIP insertion at low cell surface pH, Front. Physiol. **5**, 97 (2014).

[6] Y. K. Reshetnyak, Imaging tumor acidity: pH-Low Insertion Peptide probe for optoacoustic tomography, Clin. Canc. Res. **21**, 4502 (2015).

[7] M. Anderson, A. Moshnikova, D. M. Engelman, Y. K. Reshetnyak, and O. A. Andreev, Probe for th measurement of cell surface pH in vivo and ex vivo. PNAS **113**, 8177 (2016).

[8] R. C. Adochite, A. Moshnikova, J. Golijanin, O. A. Andreev, N. V. Katenka, and Y. K. Reshetnyak, Comparative study of tumor targeting and biodistribution of

pH (Low) Insertion Peptide (pHLIP) conjugated with different dyes, Mol. Imaging Biol. **18**, 686 (2016).

[9] Y. K. Reshetnyak, A. Moshnikova, O. A. Andreev, and D. M. Engelman, Targeting acidic diseased tissues by pH-triggered membrane-associated peptide folding, Frontiers Bioeng. Biotech. **8**, 335 (2020).

[10] F. Jähnig, Thermodynamics and kinetics of protein incorporation into membranes, PNAS **80**, 3691 (1983).

[11] A.Ben-Shaul, N. Ben-Tal, and B. Honig, Statistical thermodynamic analysis of peptide and protein insertion into lipid membranes, Biophys. J. **71**, 130 (1996).

[12] N. Ben-Tal, A. Ben-Shaul, a. Nicholls, and B. Honig, Free-energy determinants of α-helix insertion into lipid bilayers, Biophys. J. **70**, 1803 (1996).

[13] A. Kessel and N. Ben-Tal, Free energy determinants of peptide association with lipid bilayers, in *Current topics in membranes*, Vol. 52, Chap. 8, p. 205, Elsevier 2002.

[14] T. A. Rapoport, V. Goder, S. U. Heinrich, and K. E. S. Matlack, Membrane-protein integration and the role of the translocation channel, Trends in Cell Biol. **14**, 568 (2004).

[15] J. U. Bowie, Solving the membrane protein folding problem, Nature **438**, 581 (2005).

[16] T. Hessa, H. Kim, K. Bihlmaier, C. Lundin, J. Boekel, H. Andersson, I. Nilsson, S. H. White, and G. von Heijne, Recognition of transmembrane helices by the endoplasmic reticulum translocon, Nature **433**, 377 (2005).

[17] K.-I. Oh, K. B. Smith-Dupont, B. N. Markiewicz, and F. Gai, Kinetics of peptide folding in lipid membranes, Biopolymers, **104**, 281 (2015).

[18] B. van den Berg, W. M. Clemons, I. Collinso, Y. Modis, E. Hartmann, S. C. Harrison, and T. A. Rapoport, X-ray structure of a protein-conducting channel, Nature **427**, 36 (2004).

[19] A. R. Osborne, T. A. Rapoport, and B. van den Berg, Protein translocation by the Sec61/SecY channel, Annu. Rev. Cell Dev. Biol. **21**, 529 (2005).

[20] S. H. White and G. von Heijne, The machinry of membrane protein assembly, Curr. Opin. Struct. Biol. **14**, 397 (2004).

[21] P. J. Simpson, B. Schwappach, H. G. Dohlman, and R. I. Isaacson, Structures of Get3, Get4, and Get5 provide new models for TA membrane protein targeting, Structure **18**, 897 (2010).

[22] S. Brambillasca, M. Yabal, M. Makarow, and N. Borgese, Unassisted translocation of large polypeptide domains across phospholipid bilayers, J. Cell. Biol. **175**, 767 (2006).

[23] R. Renthal, Helix insertion into bilayers and the evolution of membrane proteins, CMLS **67**, 1077 (2010).

[24] G. P. Sharma, Y. K. Reshetnyak, O. A. Andreev, M. Karbach, and G. Müller, Coil-helix transition of polypeptide at water-lipid interface, J. Stat. Mech. **P01034** (2015).

[25] J.-L. Popot and D. M. Engelman, Membrane protein folding and oligomerization: the two-stage model, Biochemistry **29**, 4031 (1990).

[26] Y. K. Reshetnyak, M. Segala, O. A. Andreev, and D. M. Engelman, A monomeric membrane peptide that lives in three worlds: in solution, attached to, and inserted across lipid bilayers, Biophys J **93**, 2363 (2007).

[27] O. A. Andreev, A. D. Dupuy, M. Segala, S. Sandugu, D. A. Serra, C. O. Chichester, D. M. Engelman, and Y. K. Reshetnyak, Mechanism and uses of a membrane peptide that targets tumors and other acidic tissues in vivo, Proc Natl Acad Sci USA **104**, 7893 (2007).

[28] Y. K. Reshetnyak, O. A. Andreev, M. Segala, V. S. Markin, and D. M. Engelman, Energetics of peptide (pHLIP) binding to and folding across a lipid bilayer membrane, Proc Natl Acad Sci USA **105**, 15340 (2008).

[29] O. A. Andreev, A. G. Karabadzhak, D. Weerakkody, G. O. Andreev, D. M. Engelman, and Y. K. Reshetnyak, pH (low) insertion peptide (pHLIP) inserts across a lipid bilayer as a helix and exits by a different path, PNAS **107**, 4081 (2010).

[30] A. G. Karabadzhak, D. Weerakkody, D. Wijesinghe, M. S. Thakur, D. M. Engelman, O. A. Andreev, V. S. Markin, and Y. K. Reshetnyak, Modulation of the pHLIP transmembrane helix insertion pathway, Biophys. J. **102**, 1846 (2012).

[31] D. Vila-Viçosa, T. F. D. Silva, G. Slaybaugh, Y. K. Reshetnyak, O. A. Andreev, and M. Machuqueiro, Membrane-induced pK_a shifts in wt-pHLIP and its L16H variant, J. Chem. Theory Comput. **14**, 3289 (2018).

[32] D. Weerakkody, O. A. Andreev, and Y. K. Reshetnyak, Insertion into bilayer of truncated pHLIP peptide, Biochem. Biophys. Rep. **8**, 290 (2016).

[33] R. E. Jacobs and S. H. White, The nature of the hydrophobic binding of small peptides at the bilayer interface: implications for the insertion of transbilayer helices, Biochemistry **28**, 3421 (1989).

[34] D. M. Engelman and T. A. Steitz, The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis, Cell **23**, 411 (1981).

[35] D. M. Engelman, Y. Chen, C. N. Chin, A. R. Curran, A. M. Dixon, A. D. Dupuy, A. S. Lee, U. Lehnert, E. E. Matthews, Y. K. Reshetnyak, A. Senes, and J. L. Popot, Membrane protein folding: beyond the two-stage model, FEBS Lett. **555**, 122 (2003).

[36] G. A. Nemnes and D.-V. Anghel, Fractional exclusion statistics in non-homogeneous interacting particle systems, Roman. Rep. Phys. **66**, 336 (2014).

[37] C. Moore, D. Liu, B. Ballnus, M. Karbach, and G. Müller, Disks in narrow channel jammed by gravity and centrifuge: profiles of pressure, mass density, and entropy density, J. Stat. Mech. **P04008** (2014).

[38] B. Bakhti, P. Maass, M. Karbach, and G. Müller, Monodisperse rods in external potentials, Phys. Rev. E **92**, 042112 (2015).

[39] G. P. Sharma, S. Habeeb, M. Karbach, and G. Müller, Kinetic model for peptide insertion into a membrane, (unpublished).

[40] T. Narayanan, D. Weerakkody, A. G. Karabadzhak, M. Anderson, O. A. Andreev, and Y. K. Reshetnyak, pHLIP peptide interaction with a membrane monitored by SAXS, J. Phys. Chem. B **120**, 11484 (2016).

[41] S. C. Hanz, N. S. Shu, J. Qian, N. Christman, P. Kranz, M. An, C. Grewer, and W. Qiang, Protonation-driven membrane insertion of a pH-low insertion peptide, Angew. Chem. **128**, 12564 (2016).

[42] J. Tang, H. Yin, J. Qiu, M. J. Tucker, W. F. DeGrado, and F. Gai, Using fluorescent probes to dissect the binding, insertion, and dimerization kinetics of a model membrane peptide, J. Am. Chem. Soc. **131**, 3816 (2009).

[43] Y. K. Reshetnyak, Y. Koshevnik, and E. A. Burstein, Decomposition of protein tryptophan fluorescence spec-

tra into og-normal components III: correlation between fluorescence and microenvironment parameters of individual tryptophan residues, Biophys. J. **81**, 1735 (2001).

[44] Experiments that use a variant with the Pro residue (near center) and the Ile residue (near N terminus) switched as a control sample will shed light on this issue.

[45] G. von Heijne, Proline kinks in transmembrane α-helices, J. Mol. Biol. **218**, 499 (1991).

[46] F. S. Cordes, J. N. Bright, and M. S. P. Sansom, Prolin-induced distortions of transmembrane helices, J. Mol. Biol. **323**, 951 (2002).

[47] K. A. Williams and C. M. Deber, Proline residues in transmembrane helices: structural and dynamical role?, Biochemistry **30**, 8919 (1991).

[48] M. Tajparast and M. I. Glavinovic, Strain, stress and energy in lipid bilayer induced by electro-static/electrokinetic forces, Biochim. Biophys. Acta **1818**, 829 (2012).

[49] L. Wang, Measurements and implications of the membrane dipole potential, Annu. Rev. Biochem. **81**, 615 (2012).

[50] A. Ben-Shaul, Molecular theory of chain packing, elasticity and lipid-protein interaction in lipid bilayers, Handbook of Biological Physics, Vol. 1, Chap. 7, p. 359, Elsevier, 1995.

[51] J. Seelig, Thermodynamics of lipid-peptide interactions, Biochim. Biophys. Acta **1666**, 40 (2004).

[52] H. Khandelia, J. H. Ipsen, and O. G. Mouritsen, The impact of peptides in lipid membranes, Biochim. Biophys. Acta **1778**, 1528 (2008).

[53] L. B. Li, I. Vorobyov, and T. W. Allen, The role of membrane thickness in charged protein-lipid interactions, Biochim. Biophys. Acta **1818**, 135 (2012).

[54] K. Åman, E. Lindahl, O. Edholm, P. Håkansson, and P.-O. Westlund, Structure and dynamics of interfacial water in an Lalpha phase lipid bilayer from molecular dynamics simulations, Biophys. J. **84**, 102 (2003).

[55] J. L. MacCallum, W. F. D. Bennett, and D. P. Tieleman, Distribution of amino acids in a lipid bilayer from computer simulations, Biophys. J. **94**, 3393 (2008).

[56] The values for x_a, x_b, and x_s depend on the type of lipids used in the experiments. Some experiments use a range of lipids with different parameter values for comparative purposes.

[57] We use the name 'link' for the unit of peptide backbone between successive peptide bonds as used in the mathematical modelling of Ref. [24].

[58] A. V. Finkelstein and O. P. Ptitsyn, *Protein physics*, Academic Press, New York 2002.

[59] A. Y. Grosberg and A. R. Khokhlov, *Statistical physics of macromolecules*, AIP Press, New York 1994.

[60] D. Marsh, Lateral pressure in membranes, Biochim. Biophys. Acta **1286**, 183 (1996).

[61] R. S. Cantor, Lateral pressure in cell membranes: a mechanism for modulation of protein function, J. Phys. Chem. B **101**, 1723 (1997).

[62] R. S. Cantor, Lipid composition and the lateral pressure profile in bilayers, Biophys. J. **76**, 2625 (1999).

[63] J. Gullingsrud and K. Schulten, Lipid bilayer pressure profiles and mechanosensitive channel gaiting, Biophys. J. **86**, 3496 (2004).

[64] J. Gullingsrud, A. Babakhani, and J. A. McCammon, Computational investigation of pressure profiles in lipid bilayers with embedded proteins, Molec. Simul. **32**, 831

[65] (2006).

[65] D. Marsh, Lateral pressure profile, spontaneous curvature frustration, and the incorporation and conformation of protein membranes, Biophys. J. **93**, 3884 (2007).

[66] S. Baoukina, S. J. Marrink, and D. P. Tielemann, Lateral pressure profiles in lipid monolayers, Faraday Discuss. **144**, 393 (2010).

[67] H. Zhan and T. Lazaridis, Inclusion of lateral pressure/curvature stress effects in implicit membrane models, Biophys. J. **104**, 643 (2013).

[68] L. Dominguez, L. Foster, J. E. Straub, and D. Thirumalai, Impact of membrane lipid composition on the structure ans stability of the transmembrane domain of amyloid precursor protein, PNAS **113**, E5281 (2016).

[69] A. A. Drozdova and S. I. Mukhin, Lateral pressure profile in bent lipid membranes and mechanosensitive channels gating, bioRxiv http://dx.doi.org/10.1101/062463 (2016).

[70] This estimate compares the latent heat of sublimation, $L_{sg} = 680\text{cal/g}$ to the latent heat of melting, $L_{sl} = 80\text{cal/g}$. If both latent heats are dominated by the breaking of H-bonds, then the fraction of unsatisfied H-bonds in the liquid phase beomes $80/680 \simeq 0.12$. That percentage is even lower if we assume that the average bond strength is weaker in the liquid phase.

[71] S.-Y. Sheu, D.-Y. Yang, H. L. Selzle, and E. W. Schlag, Energetics of hydrogen bonds in peptides, PNAS **100**, 12683 (2003).

[72] P. J. Fleming and G. D. Rose, Do all backbone polar groups in proteins form hydrogen bonds? Potein Sci. **14**, 1911 (2005).

[73] D. Laage and J. T. Hynes, A molecular jump mechanism of water reorientation, Science **311**, 832 (2006).

[74] N. H. Joh, A. Min, S. Faham, J. P. Whitelegge, D. Yang, V. L. Woods Jr, and J. U. Bowie, Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins, Nature **453**, 1266 (2008).

[75] J. Gao, D. A. Bosco, E. T. Powers, and J. W. Kelly, Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins, Nature Struct. Mol. Biol. **16**, 684 (2009).

[76] Z. Cao, J. M. Hutchison, C. R. Sanders, and J. U. Bowie, Backbone hydrogen bond strength can vary widely in transmembrane helices, J. Am. Chem. Soc. **139**, 10742 (2017).

[77] We keep setting $\epsilon_h = 2\epsilon_t$ for the activation energies of hybrids ad tags, respectively as in Ref. [24] for the same reason.

[78] It is fair to argue that the coil becomes more constrained in the lipid environment than in the aqueous environment, which suggest that we should use a larger value of μ in the latter than in the former. To implement the restricted space available to the polypeptide coil in this way is technically difficult. It would entail approximations of little control and obscure other environmental effects. We, therefore, opt to account for the space restrictions imposed by the lipids in different ways.

[79] More precisely, the deviation of enthalpy density from its chosen reference value.

[80] S. Dorairaj and T. W. Allen, On the thermodynamic stability of a charged arginine side chain in a transmembrane helix, PNAS **104**, 4943 (2007).

[81] T. W. Allen, Modeling charged protein side chains in

lipid membranes, J. Gen. Physiol. **130**, 237 (2007).

[82] L. Li, I. Vorobyov, and T. W. Allen, The different interactions of lysine and arginine side chains with lipid membranes, J. Phys. Chem B **117**, 11906 (2013).

[83] A. Denath, B. Mukherjee, K. G. Ayappa, P. K. Maiti, and S.-T. Lin, entropy and dynamics of water in hydration layers of a bilayer, J. Chem Phys. **133**, 174704 (2010).

[84] G. Némethy, S. J. Leach, and H. A. Scheraga, The influence of amino acid side chains on the free energy of helix-coil transition, J. Phys. Chem. **70**, 998 (1966).

[85] W. C. Wimley and S. H. White, Experimentally determined hydrophobicity scale for proteins at membrane interfaces, Nature Struct. Biol. **3**, 842 (1996).

[86] W. C. Wimley T. P. Creamer, and S. H. White, Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides, Biochemistry **35**, 5109 (1996).

[87] S. H. White and W. C. Wimley, Hydrophobic interactions of peptides with membrane interfaces, Biochim. Biophys. Acta **1376**, 339 (1998).

[88] S. H. White and W. C. Wimley, Membrane protein folding and stability: physical principles, Annu. Rev. Biophys. Biomol. Struc. **28**, 319 (1999).

[89] H. R. Guy, Amino acid side-chain partition energis and distribution of residues in soluble proteins, Biophys. J. **47**, 61 (1985).

[90] K. Hristova and W. C. Wimley, A look at arginine in membranes, J. Membr. Biol. **239**, 49 (2011).

[91] C. P. Moon and K. G. Fleming, side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers, PNAS **108**, 10174 (2011).

[92] M. Musial-Siwek, A. Karabadzhak, O. A. Andreev, Y. K. Reshetnyak, and D. M. Engelman, Tuning the insertion properties of pHLIP, Biochim. Biophys. Acta **1798**, 1041 (2010).

[93] M. Milik and J. Skolnick, Spontaneous insertion of polypeptide chains into embranes: a Monte Carlo model, PNAS **89**, 9391 (1992).

[94] M. Milik and J. Skolnick, Insertion of peptide chains into lipid membranes: an off-lattice Monte Carlo dynamics model, Proteins **15**, 10 (1993).

[95] W. Im and C. L. Brooks, Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations, PNAS **102**, 6771 (2005).

[96] H. Nymeyer, T. B. Woolf, and A. E. Garcia, Flding is not required for bilayer insertion: replica exchange simulations of an alpha-helical peptide with an explicit lipid bilayer, Proteins **59**, 783 (2005).

[97] M. B. Ulmschneider and J. P. Ulmschneider, Membrane adsorption, folding, insertion, and translocation of synthetic transmembrane peptides, Mol. Membr. Biol. **25**, 245 (2998).

[98] M. B. Ulmschneider, J. C. Smith, and J. P. Ulmschneider, Peptide partitioning properties from direct insertion studies, Biophys. J. **98**, L60 (2010).

[99] M. B. Ulmschneider, J. P. Doux, J. A. Killian, J. C. Smith, and and J. P. Ulmschneider, Mechanism and kinetics of peptide partitioning into membranes from all-atom simulations of themostable peptides, J. Am. Chem. Soc. **132**, 3452 (2010).

[100] Y. Deng, Z. Qian, Y. Luo, Y. Zhang, Y. Mu, and G. Wei, Membrane binding and insertion of a pHLIP peptide studied by all-atom molecular dynamics simulation, Int.

J. Mol. Sci. **14**, 14532 (2013).

[101] A. A. Gurtovenko and I. Vattulainen, Calculation of the electrostatic potential of lipid bilayers from molecular dynamics simulations: methodological issues, J. Chem. Phys. **130**, 215107 (2009).

[102] G. Basu and A. Kuki, Conformational preferences of oligopeptides rich in αaminoisobutyric acid II: a model for the $3_{10}/\alpha$-helix transition with composition and sequence sensitivity, Biopolymers **32**, 61 (1992).

[103] G. L. Millhauser, Views of helical peptides: a proposal for the position of 3_{10}-helix along the thermodynamic folding pathway, Biochemistry **34**, 3873 (1995).

[104] J. A. Killian, Hydrophobc mismatch between proteins and lipids in membranes, Biochim. Biophys. Acta **1376**, 401 (1998).

[105] M. R. R. de Planque and J. A. Killian, Protein-lipid interactions studied with designed transmembrane peptides: role of hydrophobic matching and interfacial anchoring, Molec. Membr. Biol. **20**, 271 (2003).

[106] E. Strandberg, S. Esteban-Martin, A. S. Ulrich, and J. Salgado, Hydrophobic mismatch of mobile transmembrane helices: merging theory and experiments, Biochim. Biophys. Acta **1818**, 1242 (2012).

[107] A. J. de Jesus and T. W. Allen, The determinants of hydrophobic mismatch response for transmembrane helices, Biochim. Biophys. Acta **1828**, 851 (2013).

[108] The energetics of the competing α and 3_{10} helix conformations is a topic well within the reach of the methodology employed here as a natural extension of Ref. [24] on polypeptides and Ref. [131] on DNA under tension.

[109] C. Chipot, M. L. Klein, and M. Tarek, Modeling lipid membranes, in *Handbook of materials modeling*, p. 929, Springer, Dordrecht 2005.

[110] S. Ludtke,K. He, and H. Huang, Membrane thinning caused by magainin 2, Biochemistry **34**, 16764 (1995).

[111] Y. Wu, K. He, S. J. Ludtke, and H. W. Huang, X-ray diffraction study of lipid bilayer membranes interacting with amphiphilic helical peptides: diphytanyl phosphatidylcholine witj alamethicin at low concentrations, Biophys. J. **68**, 2361 (1995).

[112] R. J. Alsop, R. M. Schober, and M. C. Rheinstädter, Swelling of phospholipid membranes by divalent metal ions depends on the location of the ions in the bilayers, Soft Matter **12**, 6737 (2016).

[113] L. H. Klausen, T. Fuhs, and M. Dong, Mapping surface charge density of lipid bilayers by quantitative surface conductivity microscopy, Nature Commun. **7**, 12447 (2016).

[114] R. Friedman, Membrane-ion interactions, J. Membrane Biol. **251**, 453 (2018).

[115] I. Dobrzynska, Association equilibria of divalent ions on the surface of liposomes formed from phosphatidylcholine, Eur. Phys. J. E **4**, 3 (2019).

[116] E. Deplazes, J. White, C. Murphy, C. G. Cranfield, and A. Garcia, Competing for the same space: protons and alkali ions at the interface of phospholipid bilayers, Biophys. Rev. **11**, 483 (2019).

[117] E. Deplazes, B. D. Tafalla, C. G. Cranfield, and A. Garcia, Role of ion-phospholipid interactions in zwitterionic phospholipid bilayer ion permeation, J. Phys. Chem. Lett. **11**, 6353 (2020).

[118] S. I. Hossain, S. C. Saha, and E. Deplazes, Phenolic compounds alter the ion permeability of phospholipid bilayers via specific lipid interactions, Phys. Chem.

Chem. Phys. **23**, 22352 (2021).

[119] W. Xian, M. R. Hennefarth, M. W. Lee, T. Do, E. Y. Lee, A. N. Alexandrova, and G. C. L. Wong, Histidine-mediated ion specific effects enable salt tolerance of a pore-forming marine anti microbial peptide, Angew. Chem. Int. Ed. **e202108501** (2022).

[120] F. D. M. Haldane, Fractional statistics in arbitrary dimensions: a generalization of the Pauli principle, Phys. Rev. Lett. **67**, 937 (1991).

[121] Y.-S. Wu, Statistical distribution for generalized ideal gas of fractional-statistics particles, Phys. Rev. Lett. **73**, 922 (1994).

[122] S. B. Isakov, Statistical mechanics for a class of quantum statistics, Phys. Rev. Lett. **73**, 2150 (1994); Generalization of statistics for several species of identical particles, Mod. Phys. Lett. B **8**, 319 (1994).

[123] D.-V. Anghel, The thermodynamic limit for fractional exclusion statistics, J. Phys. A **40**, F1013 (2007); The fractional exclusion statistics amended, Europhys. Lett. **87**, 60009 (2009).

[124] G. G. Potter, G. Müller, and M. Karbach, Thermodynamics of ideal quantum gas with fractional statistics in D dimensions, Phys. Rev. E **75**, 061120 (2007); Thermodynamics of statistically interacting quantum gas in D dimensions, **76**, 061112 (2007).

[125] P. Lu, J. Vanasse, C. Piecuch, M. Karbach, and G. Müller, Statistically interacting quasiparticles in Ising chains, J. Phys. A **41**, 265003 (2008).

[126] D. Liu, P. Lu, G. Müller, and M. Karbach, Taxonomy of particles in Ising spin chains, Phys. Rev. E **84**, 021136 (2011).

[127] P. Lu, D. Liu, G. Müller, and M. Karbach, Interlinking motifs and entropy landscapes of statistically interacting particles, Condens. Matter Phys. **15**, 13001 (2012).

[128] D. Liu, J. Vanasse, G. Müller, and M. Karbach, Generalized Pauli principle for particles with distinguishable traits, Phys. Rev. E **85**, 011144 (2012).

[129] N. Gundlach, M. Karbach, D. Liu, and G. Müller, Jammed disks in a narrow channel: criticality and ordering tendencies, J. Stat. Mech. **P04018** (2013).

[130] B. Bakhti, M. Karbach, P. Maass, M. Mokim, and G. Müller, Statistically interacting vacancy particles, Phys. Rev. E **89**, 012137 (2014).

[131] A. C. Meyer, Y. Öz, N. Gundlach, M. Karbach, P. Lu, and G. Müller Molecular chains under tension: Thermal and mechanical activation of statistically interacting extension and contraction particles, Phys. Rev. E **101**, 022504 (2020).