Interactions of Phospholipid Bilayers with Several Classes of Amphiphilic alpha-Helical Peptides: Insights from Coarse-Grained Molecular Dynamics Simulations

Interactions of phospholipid bilayers with several classes of amphiphilic α-helical peptides: insights from coarse-grained molecular dynamics simulations

Paraskevi Gkeka, and Lev Sarkisov*

Institute of Materials and Processes, School of Engineering and Electronics, University of Edinburgh, UK

E-mail: Lev.Sarkisov@ed.ac.uk

Abstract

In this article, we focus on several types of interactions between lipid membranes and α-helical peptides, based on the distribution of hydrophobic and hydrophilic residues along the helix. We employ a recently proposed coarse grained model MARTINI and test the ability of the model to capture diverse types of behavior. MARTINI provides useful insights on the formation of barrel-stave and toroidal pores and on the relation between these two mechanisms. Amphipathic non-spanning peptides are also described with sufficient accuracy. The picture is not as clear for fusion and transmembrane peptides. For each class of peptides we calculate the potential of mean force (PMF) for peptide translocation across the lipid bilayer and demonstrate that each class has a distinct shape of PMF. The reliability of these calculations, as well as wider implications of the results, are discussed.

Introduction

Peptide-membrane interactions play an important role in a number of biological processes, such as antimicrobial defence mechanisms, viral translocation, membrane fusion and functions of membrane proteins. Naturally, there is a significant on-going theoretical and experimental effort to understand and elucidate these interactions.

Peptide-membrane interactions are complex and beautifully diverse phenomena. A membrane, interacting with a peptide, may experience a number of possible structural transitions, including stretching, reorganization of lipid molecules, formation of defects, transient and stable pores, vesicles and so on. In principle, computer simulation is a powerful research tool to study peptide-membrane interactions, as it is able to provide a detailed description of these processes on a molecular level. However, a model operating on an appropriate time and length scale is imperative in this description. For example, self-assembly of a trans-membrane pore from several peptide molecules may take microseconds to complete. Recently, several atomistic molecular simulation studies attempted to address peptide-membrane interactions in their full complexity. For example, Leontiadou and co-workers captured toroidal pore formation in simulations of antimicrobial peptide magainin-H2 and a model phospholipid membrane.[1] Studies of toroidal pore formation and their structural characteristics have been further extended by Sengupta and co-workers.[2] In another example, Herce and Garcia applied fully atomistic simulations to propose a complex multistage mechanism of HIV-1 TAT peptide translocation across the membrane.[3] Formation of a transient pore was observed, with the peptides diffusing on the surface of the pore to cross the membrane. An alternative mechanism, based on micropinocytosis, has been suggested for TAT translocation in fully atomistic studies by Yesylevskyy and co-workers. In micropinocytosis a cluster of peptides wraps the membrane around itself to form a small vesicle.[4] A similar mechanism of translocation was reported by the same group for another cell-penetrating peptide, Penetratin. None of these simulations however spanned timescale beyond several hundred of nanoseconds, and in many cases the runs were limited to tens of nanoseconds. Routine operation on longer time scale still remains prohibitively expensive in atomistic simulations. This limitation imposed by atomistic simulations

led to the development of coarse-grained approaches to study complex biomolecular phenomena.

Coarse-grained approaches are based on the idea of systematically reducing the level of detail in the way the system is represented, and thus increasing the time/length scale of the simulation. One way of doing this is by modeling the system as a group of effective particles ('beads'). Each of these beads represents an ensemble of atoms whose atomistic degrees of freedom do not play an important role in the process under consideration and are integrated out. This leads to several implications. First of all, it results in the expected improvement in computational efficiency of the model due to the reduced number of degrees of freedom (depending on the level of coarse-graining). Furthermore, as has been noted in a number of studies, smoothing out of fine-grained degrees of freedom in CG models reduces the effective friction between the molecules. As a result, many complex processes such as biomolecular self-assembly occur on a shorter effective time scale. This peculiar feature of GC models often provides the only opportunity to observe a complete self-assembly evolution within reasonable simulation time. Finally, the computational efficiency of CG models makes it possible to routinely calculate many important properties of biomolecular systems, that are usually inaccessible or difficult to obtain in atomistic simulations. One of these properties is the Potential of Mean Force (PMF) which provides an estimate of the free energy change along the chosen reaction coordinate. In the simulations of peptide-membrane interactions, this reaction coordinate is simply the distance between the peptide and the center of the lipid bilayer. The free energy profile along the bilayer normal provides a direct measure of the energy required to transfer the peptide between different regions of the system, such as bulk water phase, water-bilayer interface and bilayer core.

Several strategies to construct CG models have been offered over the years. For example, the interactions between coarse-grained beads can be calibrated to reproduce the forces between the corresponding groups of atoms in atomistic simulations.[5] Alternatively, the coarse-grained model can be calibrated to reproduce certain physical characteristics of the system of interest, such as density, phase transitions and structure.[6] For an excellent review of the current developments and achievements in this field we refer the reader to the article by Venturoli and co-workers.[7] State-of-

the-art in atomistic and CG simulation studies of lipid membranes, including peptide-membrane interactions, has also been recently reviewed by Marrink et al. [8]. Another recent review on the advances in the area of multiscale modeling is the one by Murtola et al. [9]

One of these models was introduced by Marrink and co-workers a few years ago, to describe properties of lipid-water systems. [6] In their model every four heavy atoms (i.e. not hydrogens) are represented by one effective bead. Four major types of bead and several variants were introduced to describe different levels of polarity and charge. For example, within this model a molecule of butane would be represented as a single apolar bead, whereas four water molecules are represented as a single polar bead. The model has been validated against several processes, such as lipid phase transitions, micellar and vesicle behavior as well as lipid bilayer formation, clearly demonstrating that such complex processes are within its scope. [10–12]

Several attempts have been made to extend the original model of Marrink and co-workers to proteins, peptides and other biological entities. One of these extended models was recently proposed by Bond and Sansom. [13,14] They introduced a model for proteins, where a representation for each amino acid was based on its properties (tendency to form hydrogen bonds, hydrophobicity/hydrophilicity and charge). The same level of coarse graining was followed (i.e. about four atoms represented as one effective bead), with the amino acids modeled by one, two or three beads, one representing the backbone of the amino acid and the others the side chain. For more information about the model, we refer the reader to the original publication. [13] Within this protocol, the authors investigated different peptides in lipid bilayers, capturing, among other effects, the insertion and dimerization of Glycophorin A (GpA), the insertion of Escherichia coli outer membrane protein OmpA and WALPs into a lipid bilayer and the interfacial orientation of a monomeric LS3 peptide. [13,14] The model has also been used for the prediction of several processes such as the insertion of DNA in a lipid bilayer, the interaction of membrane enzymes with lipid bilayers and the dependence of peptide-membrane interactions on the initial structure of the peptide in different lipid environments. [15–17]

Recently, a new version of the force field proposed by Marrink and co-workers has been de-

veloped, with an extension to proteins.[18,19] The proposed model, MARTINI, features a larger number of bead types and interactions, and has been optimized to reproduce some key properties of amino acids, such as oil/water partition coefficients and association constants between different amino acids. Moreover, the model has been shown to accurately capture peptide-membrane interactions for several helical peptides,[18] and to correctly reproduce the formation of a toroidal pore by magainin-H2, confirming earlier atomistic simulations.[1] Other applications of MARTINI include the effect of temperature and membrane composition on the properties of liposomes in the limit of high curvature,[20] the self assembly of cyclic peptides near or within membranes[21] and the formation of a barrel-stave pore by LS3 synthetic peptide.[22]

In this study, we aim to test the scope of the MARTINI model and its applicability to a wider range of systems and modes of interactions between α-helical peptides and lipid bilayers, adopted here as simplified models of lipid membranes. In this process, we will be guided by a general classification of possible modes of peptide-membrane interactions employed in a series of works by Brasseur and co-workers.[23–25] This classification is based on a view of α-helical peptides as amphiphilic entities with a well defined geometry. The idea is that the distribution of hydrophobic and hydrophilic residues along the α-helix plays a central role in the partition of the peptide between the hydrophilic aqueous media and the hydrophobic core of the lipid membrane. Thus, depending on this distribution, several possible scenarios can be identified and are schematically depicted in Figure 1. In these schematics α-helices are represented as cylinders with their hydrophobic regions shaded blue. The top part of the figure provides a side-view of the cylinders. From this side-view, one can observe the difference in the hydrophobicity distribution among various classes of peptides. At the bottom of the figure, we present the proposed interaction mechanism for each class. The helices are represented as cylinders colored orange. Although this is a simplified description, and not all of the α-helical peptides feature a well-defined distribution of hydrophobic groups, there are many α-helical peptides whose structure and behavior does indeed fall in one of these general classes. Let us briefly review some of the examples here.

Typical representatives of Type I class are the α-helical peptides forming (and derived from)

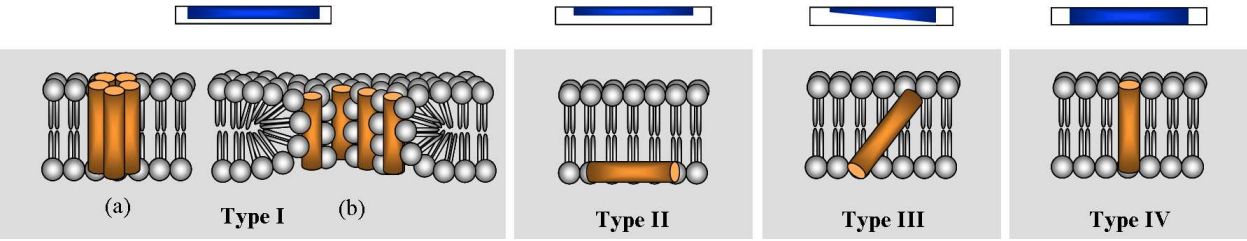


Figure 1: Schematic view of the different classes of α-helical peptides according to their hydrophobicity distribution along the axis. Top: side-view of the α-helices. The helices are represented as cylinders with their hydrophobic regions blue. Bottom: proposed interaction mechanism for each class. The helices are represented as cylinders colored orange. (Adapted from figures 3, 4 and 5 in reference 23.)

protein ion-conducting channels (Figure 1, Type I). These helices usually consist of two strands of hydrophobic and hydrophilic residues along their helical axis, with the hydrophobic region being dominant. In the membrane, several peptides form a bundle, with the hydrophobic groups facing the core of the membrane. During the past two decades, a number of model channel systems, consisting of synthetic helices of this type, have been studied. The 'synporins', synthetic peptides developed by Montal and co-workers,[26–28] as well as the 'template-assembled synthetic proteins'(TASPs) used by Mutter and co-workers,[29,30] are two examples of ion channels formed by synthetic peptides in lipid bilayers. Moreover, Lear et al. synthesized three model peptides containing only leucine and serine residues, in order to investigate the mechanism by which these helices associate into transmembrane bundles.[31] For example, one of these synthetic peptides, LS3, features all the characteristics of a membrane spanning helix as well as the necessary amphiphilicity to provide the desired aggregation of polar faces, thus leading to spontaneous formation of well-defined transmembrane ion channels. In the bundle formed by LS3 helices, the peptides are tightly aligned with the tails of the neighboring lipids stretching along them. This kind of bundle is often classified as a barrel-stave pore (Figure 1, Type I (a)).

The formation of transmembrane pores is also one of the commonly proposed mechanism of antimicrobial peptide action. In addition to the barrel-stave mechanism, antimicrobial peptides may also form transient toroidal pores, and we thus included this mechanism in the classification in Figure 1 as Type I (b). (We also note here that several other mechanisms of antimicrobial peptide

interactions, such as carpet mechanism, or various types of endo- and pinocytosis are governed by the peptide charge rather than distribution of hydrophobicity; they are not included in this classification and will not be considered in this study). In a toroidal pore the peptides are shorter than the thickness of the bilayer, and they impose a positive curvature strain on the bilayer, that leads to a toroidal structure formed by lipid leaflets and covered by peptide molecules (Figure 1, Type I (b)). Recent atomistic simulation studies of toroidal pore self-assembly reveal that the final structures are very disordered.[2] From this point of view, the schematics presented in Figure 1 is grossly oversimplified and is meant to highlight the curved structure of the lipid surface only. The coarse-grained MARTINI model is able to capture both the barrel-stave (for LS3 peptide)[22] and toroidal mechanisms (for magainin)[18] of pore formation. In this study, we will explore whether there is a link between these two processes.

The peptides whose hydrophobic region is either the same or smaller than the hydrophilic region do not have the ability to span the membrane. Examples of these peptides are provided by the synthetic lipid-associating peptide, LAP20 and the lipid-associating peptides of the plasma apolipoproteins, apoA-I. It has been shown that these peptides interact with the membrane such that the contact area of the helices with the aqueous phase is either comparable in size with that for the lipid phase (LAP20, Figure 1, Type II), or larger (apoA-I).[23] This leads to either an interfacial orientation of the peptide or formation of discoidal particles (included in the original Brasseur classification, but not considered here).

In the case of fusion peptides (Figure 1, Type III), there is a non-uniform distribution of hydrophobic residues along the helical axis. This characteristic has been suggested as one of the main reasons behind the oblique orientation of fusion peptides relative to the bilayer.[23–25] Moreover, several studies linked this particular mode of peptide insertion to the fusogenic activity of these peptides.[24,32–34] The fusion peptides of Simian Immunodeficiency Virus (SIV),[24,35,36] Newcastle Disease Virus (NDV)[37] and Human Immunodeficiency Virus (HIV)[35,38] were some of the first representatives of this class to be identified.

Those helices that are long enough to span the hydrophobic medium of the membrane, with

most or all of their residues being hydrophobic, tend to adopt a transmembrane position in a lipid bilayer (Figure 1, Type IV). These transmembrane helices have a uniform distribution of hydrophobic residues both around and along their helical axis. Among this kind of helices are the Glycophorin A (GpA),[13,39] the pHLIP peptide[40–43] as well as the TMX-1 synthetic peptide.[44]

Thus, the idea of this study is to select representative peptides from each class of behavior depicted in Figure 1 and probe the ability of the MARTINI model to capture these diverse scenarios. The selected peptides satisfy one or more of the following criteria: there should be well-reported experimental data on their behavior, they should be sufficiently simple (short), have a confirmed α-helical structure in the presence of the membranes and be of a certain technological importance. In the result section we will review some of the key properties of each peptide under consideration.

Methodology

Coarse grained model and simulation parameters

In this study, all species are described using the model of lipids and peptides recently proposed by Marrink and co-workers (MARTINI).[18,19] In MARTINI, every four heavy atoms (i.e. not hydrogens) are represented by one effective bead, with an exception made for ring structures. There are four main types of beads representing different levels of interaction: polar (P), apolar (C), nonpolar (N), and charged (Q). Apart from these main types, each bead is assigned a further subtype, in order to describe more accurately the overall chemical nature of the represented group of atoms. In this description, hydrogen-bonding capability and different levels of polarity are included. Each effective bead is assigned a mass of m=72 amu (four water molecules), and an effective size of σ=0.47 nm (this value is slightly smaller σ=0.43 nm for the ringlike structures). The van der Waals interactions are described using the Lennard-Jones 12-6 potential

$$u_{LJ}(r_{ij}) = 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right], \tag{1}$$

where i and j are the atoms, located at r_i and r_j respectively and $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ and $r_{ij} = \|\mathbf{r}_{ij}\|$. Ten different values of ε are available to represent possible interactions among various bead types (these values are scaled to 75% to describe interactions between particles belonging to a ring molecule). The bonds between the CG sites are described by a harmonic spring potential. For the description of angles a harmonic cosine potential is used. The force field has been validated against several key properties of different amino acids, such as oil/water partition coefficients and free energy profiles of amino acid insertion into a model lipid bilayer.[18] For a more detailed description of the force field parameters, we refer the reader to the article by Monticelli et al.[18]

The coarse grained simulations presented in this publication are performed with the GROMACS simulation package, version 3.3.2.[45] Initially, we perform molecular dynamics simulations of lipid and water components, in order to obtain the initial bilayer structures that are used in our studies. These preliminary simulations are up to 200 ns long depending on the size of

the bilayer. The protocol and the simulation parameters used here have previously been employed by Marrink and co-workers.[6] Three different lipid bilayer systems are considered: 1,2-Dipalmitoyl-sn-Glycero-3-phosphocholine or DPPC (256 lipids, 3228 waters), 1,2-Dioleoyl-sn-Glycero-3-phosphocholine or DOPC (128 lipids, 1500 waters and 512 lipids,12000 waters) and 1-Palmitoyl,2-oleoyl-sn-Glycero-3-phosphocholine or POPC (512 lipids, 6000 waters and 512 lipids,12000 waters). We choose different lipid systems in order to have a direct comparison with the corresponding experimental studies for each peptide.

The atomistic structures of the peptides are generated using HyperChem 8.0 software.[46] To coarse-grain these structures, we apply the script provided on Prof. Marrink's research group webpage.[47] All the peptides are capped at their termini apart from the LS3 peptide. After minimizing the energy of individual peptide molecules, we randomly insert them in the system of interest. In the cases where the peptides under study are charged, ions are inserted in the system to maintain the overall system electroneutrality. For each system, we then perform energy minimization using the steepest descent method. Finally, molecular dynamics (MD) simulations with constant pressure, temperature and number of particles (NPT ensemble) are performed. The temperature is kept constant for each group, at 323 K for the DPPC/peptides systems and 300 K for the DOPC/peptide and POPC/peptide systems, using the Berendsen thermostat with a relaxation time of 1 ps.[48] The pressure of the system is semi-isotropically coupled and maintained at 1 bar using the Berendsen algorithm with a time constant of 5 ps and a compressibility of 4.5×10^{-5} bar^{-1}.[48] The nonbonded potential energy functions are cut off and shifted at 12 Å, with forces smoothly decaying between 9 Å and 12 Å for van der Waals forces and throughout the whole interaction range for the treatment of electrostatic forces. The simulations are performed using a 25-fs integration time step. The simulation parameters applied in our membrane-peptide studies have previously been proposed and used by other groups.[13,18]

Atomistic Simulations

For the specific case of the SIV fusion peptide we perform atomistic molecular dynamics simulations in a DOPC bilayer. We use the united atom lipid parameters initially developed by Berger et al.,[49] and modified for DOPC lipids by Siu et al.,[50,51] combined with the GROMOS96 force field and Simple Point Charge (SPC) model for water proposed by Berendsen.[52] We assume that the peptide is in an α-helical secondary structure, apart from its C-terminus which is left flexible, in agreement with previous observations.[53] To maintain the secondary structure, we put restraints between the i^{th}-($i^{th}+4$) α-carbons, starting from the third residue of the peptide.

From a number of preliminary simulations, up to 100-ns-long, only a few resulted in the actual lipid bilayer formation. From these, we choose a system of 97 DOPC molecules and 4947 water molecules. We then randomly insert the SIV fusion peptide in the water phase of the system, and perform NPT molecular dynamics simulations. All simulations are carried out using the GROMACS simulation package,[45] following the approach proposed by de Vries et al.,[54] at $T = 300$ K, $P = 1$ bar, with a timestep of 2.5 fs, using the Berendsen thermostat and barostat.[48]

Potential of Mean Force Calculations

We are interested in the Potential of Mean Force (PMF) as a function of the distance between the peptide and the lipid bilayer. In this study the PMF is calculated using the Umbrella sampling protocol.[55] The total separation distance between the peptide and the center of the bilayer (5 nm) is divided into 50 small windows of 0.1 nm each. In each window, a 100-ns-long simulation is performed, with the biasing potential applied to restrain the center of mass of the peptide at a required distance from the center of the bilayer. Thus, a single PMF calculation requires 50 simulations, covering the whole separation range of interest, with the total simulation time of 5 μs. A force constant of 1000 kJ mol^{-1} nm^{-2} is applied, following the approach by Monticelli et al.[18] The system used in the case of the longer peptides, TMX-1 and pHLIP, features a large enough water phase to avoid possible effects associated with the system size and peptide-peptide interactions over periodic boundaries. All the peptides are left free to rotate around their restrained center of

mass. The sampling histograms overlap sufficiently showing that the helices move around into the neighboring windows. Moreover, all histograms contain roughly the same number of data points ($4 \cdot 10^6$ each). Finally, in order to obtain the unbiased PMFs, we use the weighted histogram analysis method (WHAM),[56] with 50 bins and a tolerance of 10^{-5}kT for the convergence of WHAM equations.

Results

The goal of this study is to test the scope and applicability of a recently introduced coarse-grained model, MARTINI, to peptides with different characteristics and different behavior in a lipid bilayer. Below, in Table 1, we summarize the peptides considered in this work and the class of behavior they belong to.

Table 1: Summary of the peptides under study and their primary sequences.

Peptide	**Sequence**	**Type**	**Ref.**
LS3	$(LSSLLSL)_3$	I	31
LAP20	VSSLLSSLKEYWSSLKESFS	II	57
SIV	GVFVLGFLGFLA	III	53
TMX-1	WNALAAVAAALAAVAAALAAVAASKSKSKSK	IV	44
N-pHLIP	ACEQNPIYWARYANWLFTTPLLLLNLALLVDADEGTG	IV	40

In Figure 2, we present the side and top view of the peptides under study. For simplicity, we represent the backbone beads of the helices with orange, the side chain beads of all the hydrophilic residues with light blue and all the side chain beads of the hydrophobic residues with dark blue. Alanine is represented by one bead colored dark blue as an indication of its hydrophobic nature. In the top view of the helices, the hydrophilic beads have been removed in order to show the distribution of the hydrophobic beads around the helical axis.

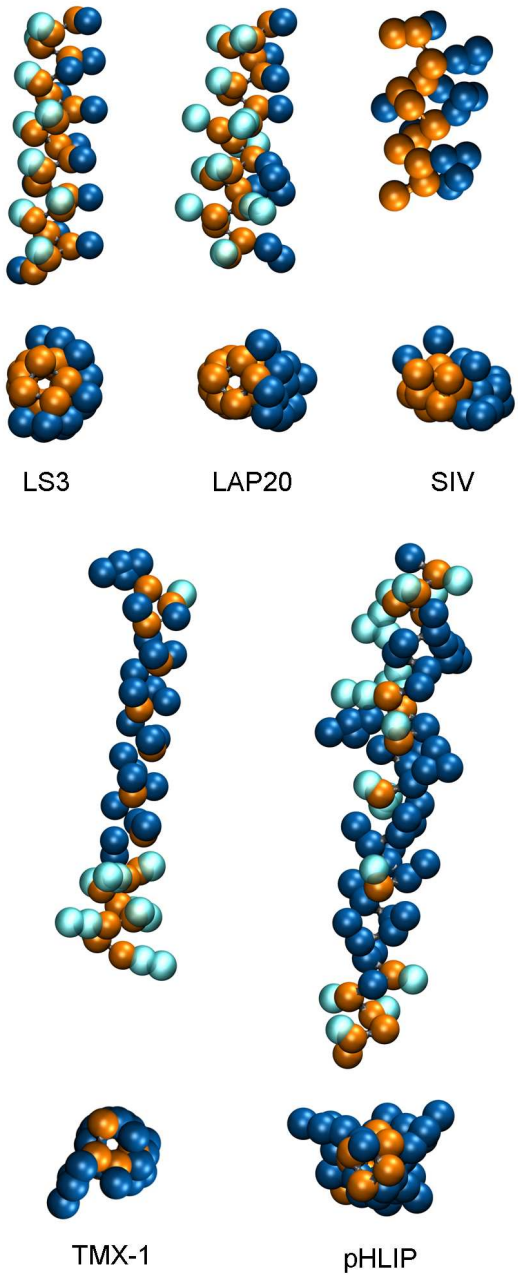


Figure 2: Side and top views of LS3, LAP20, SIV, TMX-1 and pHLIP peptides. Their backbone beads are shown in orange, their hydrophobic residues in dark blue and their hydrophilic residues in light blue. Alanine is represented by one dark blue bead.

Pore-forming peptides

The conduction of ions across a membrane is an important biological process performed by ion channel proteins. It is therefore of a great interest to understand how these structures form and function. The synthetic peptides suggested by Lear et al. are an excellent model system for this study, since they are simple, and there is experimental evidence that they form bundles with properties similar to that of the ion channel proteins.[31] One of these model peptides, LS3, is a 21-residue amphiphilic peptide with a repeating motif $(LSSLLSL)_3$ (Figure 2). Its hydrophobic residues (leucine, L) and its hydrophilic residues (serine, S) form two parallel bands on the surface of the helix, as shown in Figure 2. Due to its amphiphilic nature, LS3 shows a tendency to hide its hydrophobic residues either by adopting an interfacial orientation when monomeric, or by taking part in the formation of pores. The formation of ion conducting bundles by approximately six LS3 helices has been confirmed in a number of studies.[31,58–61] Thus, LS3 can be classified as a pore-forming peptide (Figure 1, Type I). It has also been reported that the application of transmembrane voltage significantly enhances the pore formation, due to asymmetric charge distribution within the helix (the N-terminus is positive and the C-terminus is negative).[62]

We perform a range of molecular dynamics simulations with peptide/lipid ratios from 6/256 to 12/256, in a DPPC membrane. We randomly put the peptides in a system with a preassembled DPPC lipid bilayer, and run MD simulations for several microseconds. We observe the formation of different complexes as well as the interfacial orientation of the peptides. The formation of complexes and pores seems to be initiated when some of the peptides adopt a transmembrane orientation. The formation of dimers or trimers appears to be more favorable than a transmembrane orientation of a single peptide. Regarding the orientation of the termini of the helices within the bilayer, there seems to be no particular preference in the absence of the transmembrane potential. It is also worth mentioning that at higher concentrations of the peptide, the propensity to form large complexes and pores is higher, while at smaller concentrations dimers and trimers are more common. However, a high concentration of peptides does not necessarily lead to a pore formation. In Table 2, we present a summary of the simulations performed.

Table 2: Summary of the performed simulations and key observations for LS3 peptide. In the cases where complexes or pores are formed, the remaining peptides are in interfacial orientation.

Concentration	**Duration** (μs)	**Behavior**
6 peptides	5.0	interfacial
	5.4	trimer
	6.3	interfacial
7 peptides	5.5	dimer, trimer
8 peptides	6.0	interfacial
	6.2	trimer
	17.0	2 trimers
9 peptides	5.9	interfacial
	6.0	interfacial
	6.2	trimer
10 peptides	5.8	dimer
	5.8	interfacial
	11.2	pentamer
	17.2	hexamer
	18.7	trimers/hexamer
11 peptides	5.1	tetramer
	5.6	tetramer, dimer
	5.9	tetramer
12 peptides	5.2	interfacial
	6.2	trimer/tetramer

From the observed complexes, the dimers and trimers are formed at different concentrations and are stable for several microseconds. The tetramers are also very stable (for more than 10 microseconds in some cases) with some of them having the shape of a pore. A complex of five helices has been observed once, without however resembling a pore. The hexameric bundle observed in one of our simulations can be classified as a barrel-stave pore based on its shape and structure of the lipids around it.[22] The internal diameter of the pore is about 5.2 Å, and is in good agreement

with previous experimental studies.[31,58] It is also important to note that the pore is filled with water during the whole simulation time since the formation of the pore (Figure 3). A complex consisting of six helices, without the formation of an actual pore has also been observed. In Figure 4, we present the top view of complexes of different sizes. As shown in the figure, most of the hydrophilic residues of the peptides are lying at the inner surface of the complexes while the hydrophobic particles at the outer surface.

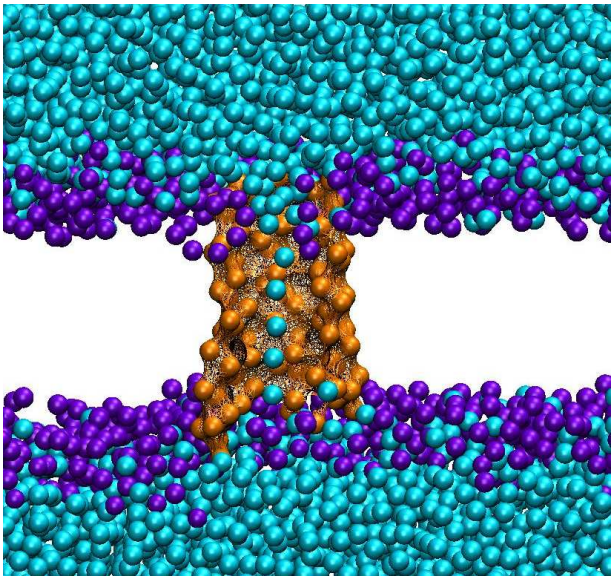


Figure 3: A barrel-stave transmembrane pore formed by six LS3 peptides. During the whole simulation time after its formation the pore is filled with water. Colors: cyan=water, orange=peptides' backbone beads, purple=phospholipid heads. For reasons of clarity the lipid tails are not shown. The beads are not to scale.

Computer visualization of the dynamics of the system where the hexameric barrel-stave pore is formed shows that the mobility of the peptides within the bundle is quite limited; however, the bundle as a whole is able to freely move within the bilayer plane. To estimate the lateral self-diffusion coefficient of the bundle within the lipid bilayer, we perform seven 360-ns-long simulations starting from different initial configurations. We then compute the average mean-square bundle displacement from the seven runs. The lateral diffusion coefficient, D_{lat}, is given by

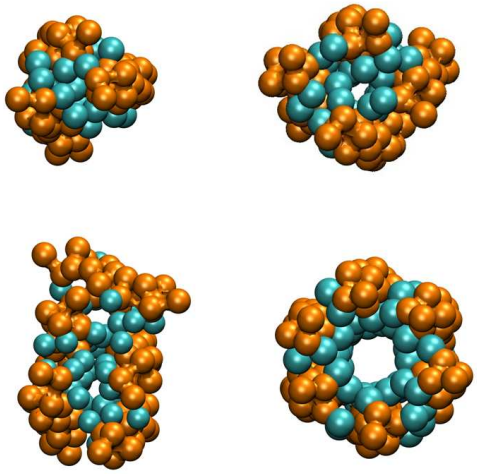


Figure 4: Top view of different types of complexes observed in LS3 simulations. The backbone beads are presented in orange and the hydrophilic residues in cyan. Other groups are omitted for clarity.

$$MSD = 4D_{lat}t. \tag{2}$$

Thus, we can extract the lateral self-diffusion coefficient by fitting a straight line to the mean-square displacement (MSD) of the bundle (Figure 5). The mean square displacement as a function of time is shown in Figure 5, and it seems that two distinct regimes of behavior can be identified. Specifically, below 100 nanoseconds the slope of the curve is clearly steeper than that for longer times. In molecular dynamics simulation, a mixture of ballistic and linear Einstein diffusion mechanisms is sometimes observed on short time scales. Thus, we omit the data up to 100 nanoseconds and fit a straight line using the data between 100 ns and 360 ns. It has been noted in previous studies that the effective dynamics appear to be faster in coarse-grained simulations because of the smoothed potentials.[19] To take this into account, we scale the calculated self-diffusion coefficient by a factor of 4, as suggested in earlier studies.[19] We finally have,

$$D_{lat} = \frac{7.3 \times 10^{-6}}{16}\ nm^2/ps$$
$$= 0.45\ \mu m^2/s.$$

Remarkably, this value is close to reported values for membrane proteins. For example, for tetraspanin CD9, that consists of 229 amino acids (LS3 pore consists of 126 amino acids), the lateral diffusion coefficient was calculated to be 0.23 $\mu m^2/s$,[63] whereas for bacteriorhodopsin, which is formed by seven transmembrane α-helices, the diffusion is 0.31 $\mu m^2/s$.[64]

It is also worth mentioning that the relative angles between the helices are very small (a few degrees), as the helices are strongly aligned, whereas each of the helices adopts an orientation of around 20^o relative to the bilayer normal, resulting in an similar orientation (of 20^o) of the bundle as a whole relative to the bilayer normal, which is also in good agreement with previous studies.[65]

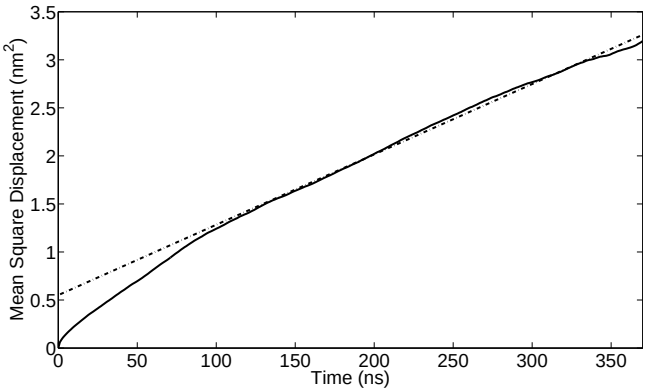


Figure 5: Mean square displacement versus time (solid line) and fitted line (dashed line). The fit is applied to the data beyond the first 100 ns and has the unscaled slope of 7.3 10^{-6} nm^2/ps.

Another notable case from our simulations is the formation of a hexameric complex when two trimers approach each other and merge into one big structure. The trimers are stable for about 10 μs and when their relative distance becomes small, a hexamer is formed. In Figure 6, we present top view snapshots from the formation of the hexamer. In the figure, the backbone beads are colored orange and water is colored blue.

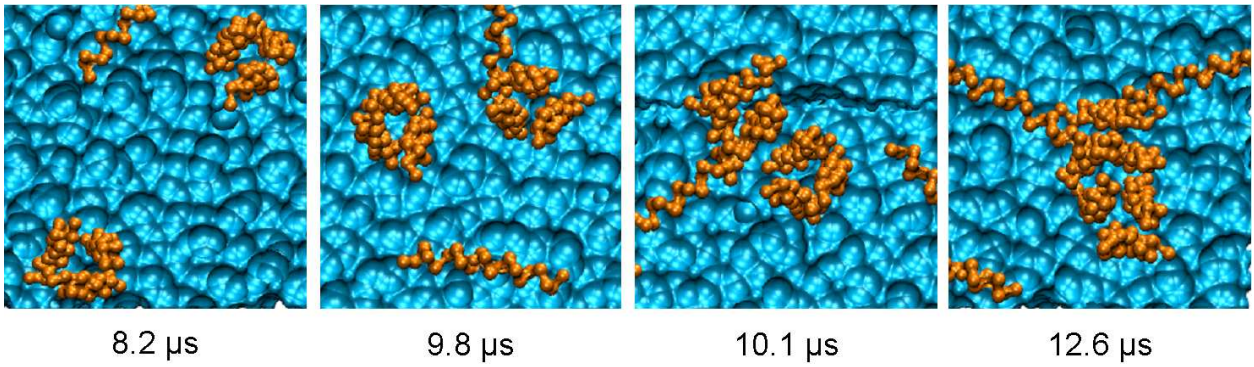


Figure 6: Snapshots from the formation of a hexameric complex from two trimers.

In Figure 7, we present the potential of mean force (PMF) for the transfer of LS3 peptide across a DPPC lipid bilayer. The PMF is represented by a continuous line whereas the dashed lines correspond to the averaged location of the phospholipid heads. The peptide seems to have two favorable positions in the lipid bilayer; one close to the lipid heads and one in the hydrophobic core of the membrane, with the first minimum most likely associated with the interfacial position of LS3 as a monomer. The first minimum (at the bilayer interface) is about -43 kTs compared to the water phase, while the minimum in the center of the bilayer is another 6kTs lower compared to the first minimum. It is worth mentioning that a similar PMF has been reported for the WALP23 peptide, with the first minimum being about -60kTs compared to the water phase and -70 kTs for the minimum in the center of the bilayer.[66]

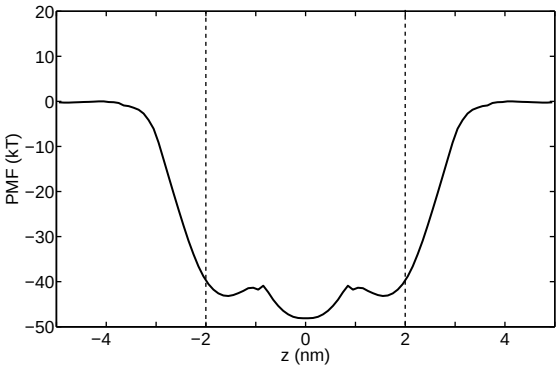


Figure 7: Potential of mean force for the transfer of the LS3 synthetic peptide from the water phase across a DPPC lipid bilayer.

$(LSSLLSL)_2$

A smaller version of LS3 has also been studied through the experiments.[58] In this peptide, the heptad of leucines and serines (LSSLLSL) is repeated twice instead of three times. In the study by Akerfeldt et al., there was no evidence for a discrete channel formation by $(LSSLLSL)_2$.[58] Here, as in the case of LS3, we perform long MD simulations (up to about 8μs) of $(LSSLLSL)_2$ in a DPPC lipid bilayer for different concentrations of the peptide (from 6 to 12 peptides in a lipid bilayer of 256 lipids). In Table 3, we present a summary of the simulations performed with $(LSSLLSL)_2$. Generally, an interfacial orientation is preferred. In some cases, when a number of the helices are initially placed in the bilayer rather than in the aqueous phase, we observe the formation of bundles of different size, but no dimers or trimers. The length of $(LSSLLSL)_2$ is less than the thickness of the bilayer. As a result, in order to position its hydrophilic residues away from the hydrophobic lipid tails, it creates complexes that involve both peptides inserted in the bilayer and located at the bilayer interface. In general, these complexes are much less ordered compared to the barrel-stave structures. A similar structure has been observed for antimicrobial peptide magainin-H2 (as well as melittin), both via atomistic simulations and CG simulations with MARTINI, and has been classified as a toroidal pore.[1,18] In an effort to compare these observations, we repeat the simulations with magainin-H2 and manage to reproduce the formation of a toroidal pore. In Figure 8, we show a snapshot from a simulation with $(LSSLLSL)_2$, with a pore consisting three inserted and three tilted helices (right) and a snapshot from simulations with magainin-H2 (left). Although the two pores have considerably different sizes, which is due to the difference in the length of the helices, one can observe that they share some common features, for example in both cases there are peptides at the bilayer/water interface that seem to stabilize the pore.

Table 3: Summary of the simulations for $(LSSLLSL)_2$ peptide. For each toroidal pore we report a pair of numbers that correspond to the fully inserted peptides in perpendicular orientation(first number) and the peptides that are inserted at oblique angle close to the bilayer-water interface(second number).

Concentration	**Duration (μs)**	**Behavior**
6 peptides	2.8	interfacial
	2.8	interfacial
	3.0	interfacial
7 peptides	2.9	interfacial
	3.2	interfacial
8 peptides	2.6	interfacial
	3.8	interfacial
9 peptides	6.0	4-6 toroidal
	8.2	4-8 toroidal
	8.2	4-8 toroidal
10 peptides	3.2	interfacial
	3.5	interfacial
	6.9	3-6 toroidal
11 peptides	3.0	interfacial
	3.3	interfacial
12 peptides	3.1	interfacial
	3.2	interfacial
	3.3	interfacial

Several variations of this toroidal self-assembly are observed in $(LSSLLSL)_2$ simulations. Typically, the complexes consist of three or four peptides in the hydrophobic area of the lipids with another three to four peptides close to the membrane/water interface, creating a toroidal-shaped structure. Because of the size of the internal diameter of the bundles, no water has been observed in these bundles. Moreover, they appear to be less stable than the ones observed for LS3. In some cases, they initially consist of three peptides inside the membrane and three at an oblique angle

close to the interface, and after some hundreds of nanoseconds their size may change (they become smaller or bigger). In one of the simulations, the pore is disassembled after about a microsecond, whereas in the case of LS3 all the formed complexes are stable for several microseconds until the end of the simulations.

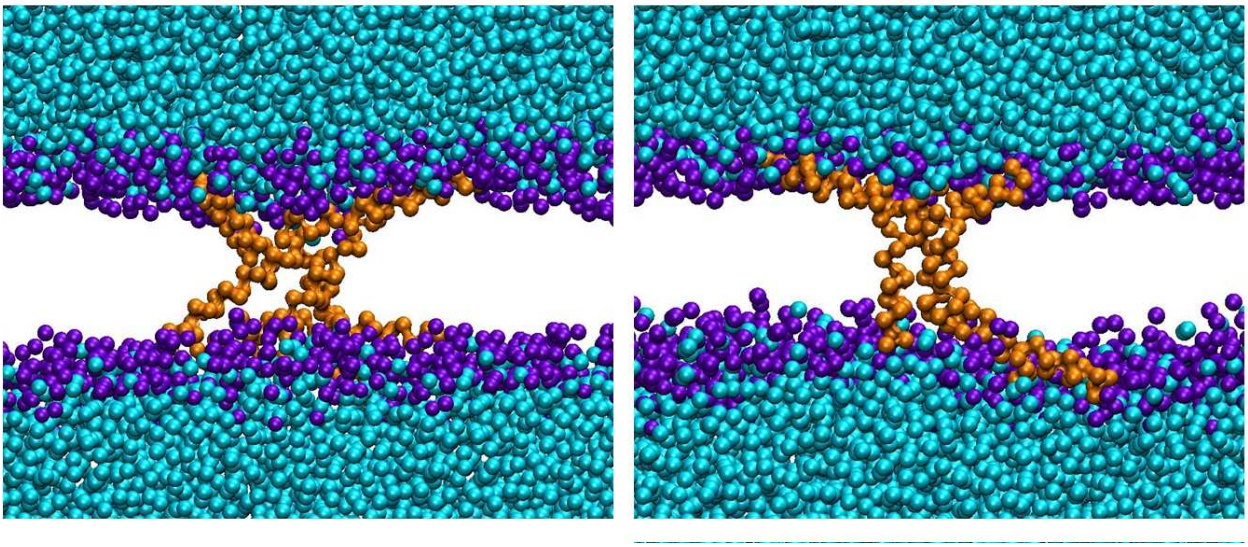


Figure 8: Snapshot of a simulation with magainin-H2 (*left*) and $(LSSLLSL)_2$ peptide (*right*) forming toroidal pores.

From the free energy calculations for $(LSSLLSL)_2$, one can observe a remarkable difference of the PMF profiles for $(LSSLLSL)_2$ and $(LSSLLSL)_3$ (Figure 9 and Figure 7). The PMF for $(LSSLLSL)_2$ features strongly pronounced minima at the interfacial positions (about -30 kTs, compared to the water phase), while the center of the bilayer is not a preferred location (+3.5 kTs compared to the water phase). This is consistent with the observations from the other molecular dynamics simulations, where $(LSSLLSL)_2$ strongly prefers the interfacial orientation, and can be positioned inside the bilayer only as part of a larger entity, with these entities being quite unstable. Thus, by changing the length of the peptide (in other words by changing the hydrophobic match), we can drastically affect the type of the observed behavior.

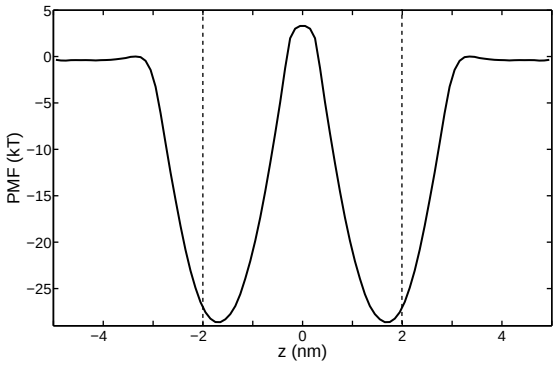


Figure 9: Potential of mean force for the transfer of $(LS3)_2$ peptide from the water phase across a DPPC lipid bilayer.

Amphipathic non-spanning helices

The second class of peptides under study is the amphipathic non-spanning helices. In this work, we focus on LAP-20 as a typical representative of this class. LAP-20 (VSSLLSSLKEYWSSLKESFS) is a synthetic lipid-associating peptide with a behavior similar to that of apolipoproteins.[57] This peptide adopts an α-helical secondary structure in the vicinity of a lipid bilayer. In Figure 2, we can see that LAP-20 has most of its hydrophobic residues grouped together on the same side along its helix. Because of this distribution of its hydrophobic and hydrophilic residues, LAP-20 adopts an interfacial orientation relative to a lipid bilayer (Figure 1, Type II).[23]

We perform MD simulations of one LAP-20 peptide in a DOPC lipid bilayer to match experimental studies. During the 1 μs of our simulations, LAP-20 adopts an interfacial orientation with most of its hydrophobic residues hidden in the phospholipid heads (Figure 10). Furthermore, we calculate the angle distribution of the peptide relative to the bilayer normal (Figure 11). For this we calculate the principal axes of inertia for the LAP20 α-helix and identify the one associated with the longer dimension of the helix. The peptide orientation is calculated as the angle between this axis and the bilayer normal, defined as a line perpendicular to the plane formed by the lipid heads. From the figure, it is evident that LAP20 has a preference for an interfacial orientation as expected from the previous studies.[23]

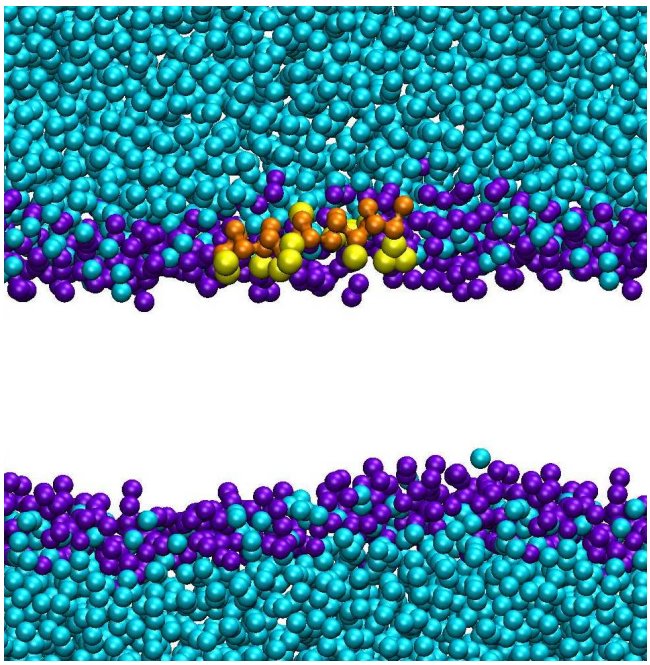


Figure 10: Snapshot from the MD simulation of the LAP20 synthetic peptide. The backbone beads of the peptide are colored orange, the hydrophobic side chains yellow, the water light blue and the phospholipid heads purple. For reasons of clarity the hydrophilic side chains and the lipid tails are not shown. The beads are not to scale.

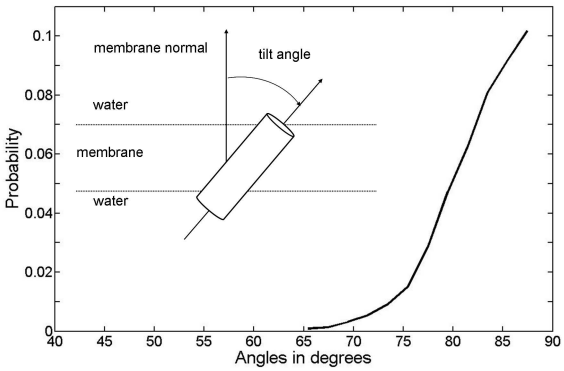


Figure 11: Angle distribution for the LAP-20 peptide. Subfigure: schematic representation of the angle definition.

The PMF profile of the LAP20 synthetic peptide is consistent with the behavior seen in the MD simulations (Figure 12). This PMF has a minimum, about -40 kTs, at the membrane/water interface in agreement with the observed location and orientation of the peptide. The center of the bilayer is an energetically unfavorable location with a maximum of more than 35 kTs compared to the water phase. Furthermore, there are clear similarities between this PMF and the PMF calculated

for $(LSSLLSL)_2$, with the features of the $LSSLLSL)_2$ PMF magnified in the LAP20's case. Thus, it seems that from this point of view the two peptides belong to the same class of non-spanning peptides. Whether or not under certain circumstances LAP20 exhibits self-assembly behavior similar to the toroidal structures of $(LSSLLSL)_2$ remains an open question.

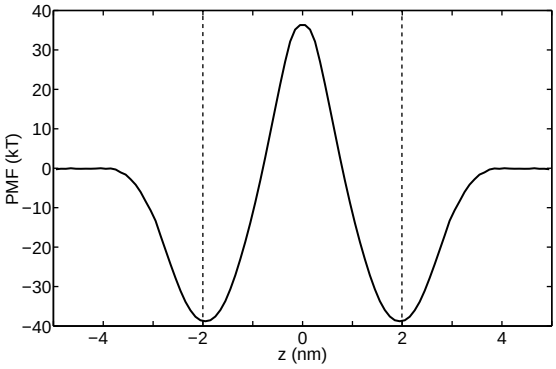


Figure 12: Potential of mean force for the transfer of LAP20 synthetic peptide from the water phase across the DOPC lipid bilayer.

Fusion peptides

In general, most of the fusion peptides share some common features: they are short, about 10 to 20 residues long and α-helical in the presence of a lipid membrane, with a gradient of hydrophobicity along their axis. It has been proposed in a number of studies that this disbalance of hydrophobicity along the α-helix leads to an oblique orientation of the peptide within the membrane and this orientation is intimately linked to the ability of the peptide to induce and facilitate membrane fusion. Efremov et al. further noted that most of the fusion peptides have access to a wide conformational space in a membrane, which also may be one of the key factors leading to the perturbation of the membrane and ultimately fusion.[67]

In our studies, we focus on one of the most studied fusion peptides, the fusion peptide of Simian Immunodeficiency Virus (SIV). This peptide features twelve residues (GVFVLGFLGFLA) and is α-helical in the presence of lipid bilayers (Figure 2). All residues of the SIV peptide are hydrophobic (except for the small, weakly hydrophilic glycine) and the disbalance in the hydrophobicity

characteristic for fusion peptides arises from the aromatic rings of three phenylalanine residues aligned and grouped together on one side of the helix. Bradshaw and co-workers have performed a series of important studies on SIV fusion peptide.[53,68,69] In one of them, the authors carried out neutron diffraction measurements from stacked multilayers of DOPC and determined the location and orientation of specifically deuterated SIV fusion peptides within the bilayer.[53] The results from this study showed that there are two different populations of peptides; one major population close to the bilayer surface, and a smaller population hidden in the hydrophobic core. Two equally plausible orientations at 55^o and 78^o with respect to the bilayer normal, were found consistent with the experimental observations. However, based on the additional FTIR data from previous studies,[70] the oblique orientation at 55^o was accepted as the most probable one.

We perform coarse-grained MD simulations with SIV fusion peptide in a DOPC lipid bilayer. To calculate the average orientation of the peptide relative to the bilayer normal, we perform an analysis of the angle distribution similar to that employed for LAP20. In order to closely reflect the experimental evidence which indicates that the C-terminus of SIV tends to be more disordered,[53] we excluded the first two residues at the C-terminus from participation in the $\alpha-$helical secondary structure. (This is achieved simply by removing secondary structure constraints imposed by MARTINI for the beads of the first two residues). The principle axes of inertia are then calculated based on the residues in the $\alpha-$helical formation only. The results are presented in Figure 13. SIV prefers to be at 70^o relative to the bilayer normal, but a wide range of angles from about 45^o to almost completely horizontal orientation is explored by the peptide. This last observation seems to be consistent with the ability of fusion peptides to access a wide range of configurations.[67,71] However, the actual preferred orientation does not seem to be in agreement with the oblique angle of 55^o suggested by Bradshaw and co-workers.[53] To test the reliability of this result we perform a fully atomistic simulation of the SIV peptide, interacting with a DOPC bilayer (see Methodology for the details of the forcefield). The orientation of the peptide is assessed using the same technique as in the coarse-grained simulations (again, the first two residues at the C-terminus do not participate in the α-helix and are not included in the angle distribution analysis). From the

results presented in Figure 13, it is clear that a similar distribution of angles is observed in atomistic simulations and thus the source of the discrepancy of these results with the experiments must lie elsewhere. Figure 14 shows a typical orientation of the SIV peptides in the atomistic and CG simulations.

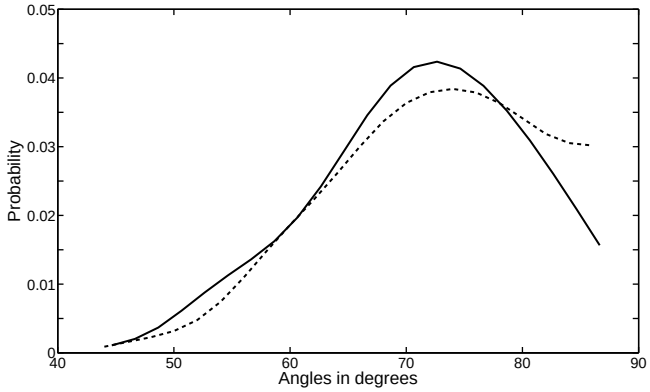


Figure 13: Angle distribution for the SIV peptide from atomistic (solid line) and coarse-grained simulations (dashed line).

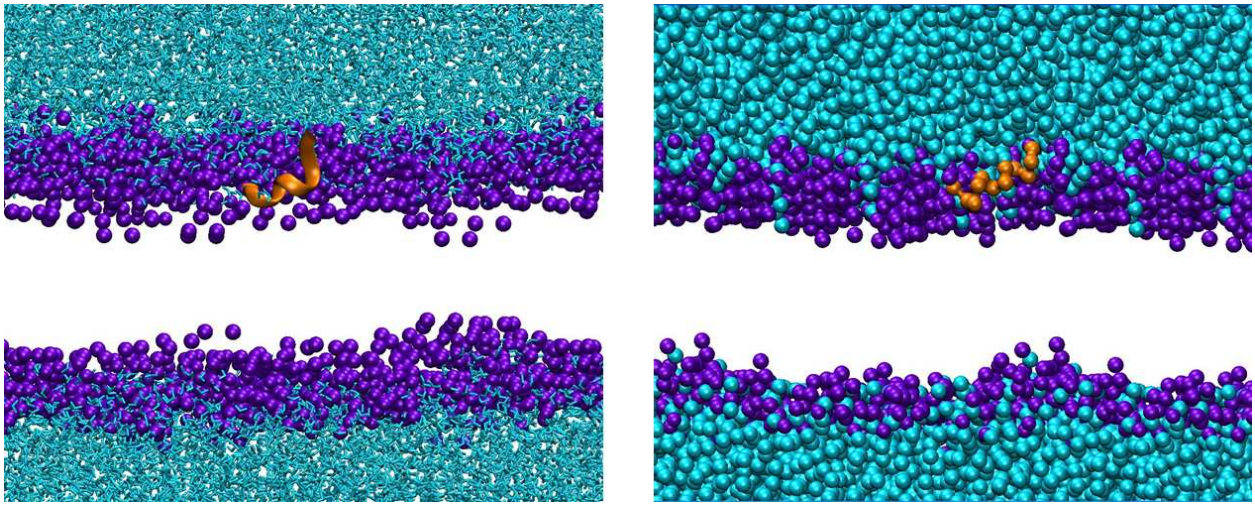


Figure 14: Snapshots from atomistic *(left)* and coarse-grained *(right)* simulations of SIV fusion peptide in a DOPC lipid bilayer. The backbone of the peptides is shown in orange, water is shown in blue and the phospholipid heads in purple. The side chains of the helices are not shown for clarity. The beads are not shown to scale.

To further extend qualitative comparison of the SIV behavior with the experimental results, we calculate the density profiles of different residues of the peptide in a lipid bilayer. In the original publication by Bradshaw and co-workers, density profiles for deuterated valine 2, leucine

8 and leucine 11 are presented and serve as the main evidence of the two possible locations of the peptide within the bilayer.[53] The double pick is particularly evident for leucine 8 and therefore we focus here on this residue. In Figure 15, we present the density profile for leucine 8 of the SIV fusion peptide from both the atomistic and the coarse grained simulations. Since the thickness of the bilayer is somewhat different in the atomistic and coarse-grained representations, the z-axis is given in dimensionless units with the lipid length in a particular representation being the scaling parameter. The atomistic simulations predict a deeper positioning of leucine 8 in the bilayer whereas in the CG simulations the peptide lies closer to the phospholipid heads. Neither of the simulations generates a double pick in density as observed in the experiments.

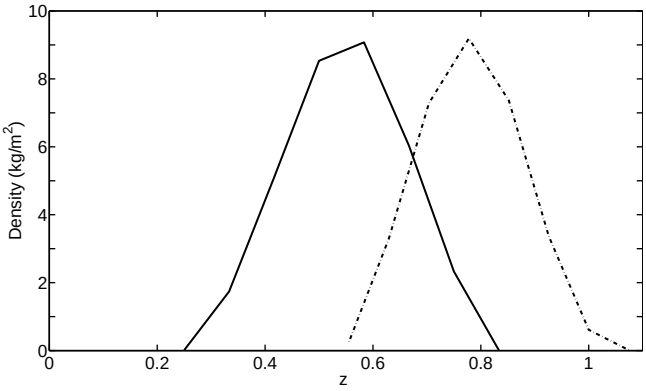


Figure 15: Density profiles of leucine 8 as a function of the distance from the center of the bilayer z. z has been normalized with the appropriate lipid length corresponding to the atomistic and the coarse-grained simulations. Solid line corresponds to the atomistic simulation and the dashed line to the coarse-grained simulation, respectively.

Finally, to complete the analysis we present the PMF calculations for the SIV peptide in Figure 16. Note, that the shape of the PMF is quite different from those observed for the other classes of peptides. We believe the PMF presented in Figure 16 is consistent with the MD behavior of the SIV peptide. Specifically, the two minima in the PMF correspond to the interfacial location, whereas the center of the bilayer is a less preferred location but not fully excluded (particularly, when compared with LAP20 or even $(LSSLLSL)_2$ peptides).

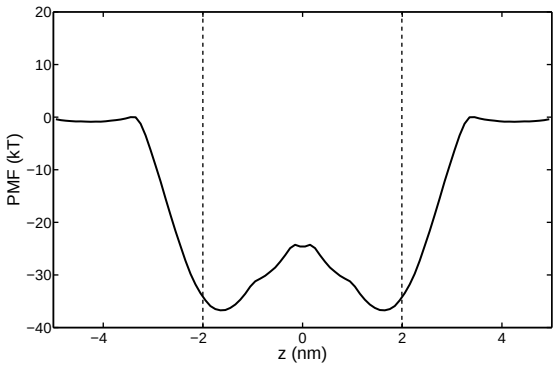


Figure 16: Potential of mean force for the transfer of the SIV fusion peptide from the water phase across a DOPC lipid bilayer.

Transmembrane helices

In the original description of various classes of peptides provided in Figure 1, transmembrane helices span the lipid bilayer due to the match between the hydrophobic region of the α-helix and the width of the hydrophobic core of the lipid bilayer. Alternatively, one could employ a functional definition of transmembrane peptides as structures capable of spontaneous insertion in the bilayer, leading to a characteristic transmembrane orientation. This ability of transmembrane peptides to cross lipid membranes (or bilayers) is of great interest in the development of novel drug vectors and other applications. However, the details of the translocation mechanisms are not yet understood. Studies of these peptides have been carried out both through experiments and simulations. For example, Bond and Sansom captured the spontaneous insertion of GpA helix by means of CG simulations.[13] Another peptide, that has been widely studied in a series of experiments by the group of Engelman, is pH (low) insertion peptide, known as pHLIP.[40–43] pHLIP (ACEQNPIYWARYADWLFTTPLLLLDLALLVDADEGTG) is a peptide derived from the integral membrane protein bacteriorhodopsin C. It exhibits high solubility at neutral pH in a nonhelical conformation, however, at lower pH (pK_a of 6.0), the peptide adopts a transmembrane position in an α-helical conformation. It has been proposed that this pH-dependent mechanism of translocation can be exploited in the early detection of pathological conditions in cells. In one of their studies, Engleman and co-workers, designed two variants of pHLIP in order to check the specificity of the function of

the peptide in tumors and test its mechanism of insertion. In one of the variants, N-pHLIP (ACE-QNPIYWARYA**N**WLFTTPLLLLL**N**LALLVDADEGTG), the Asp residues are replaced by Asn. This peptide remains an α-helical structure in the presence of liposomes over a wide pH range and adopts a transmembrane position in a lipid bilayer.

In our studies, we choose N-pHLIP as a characteristic example of a transmembrane helix and as a technologically important case. We perform MD simulations with two different systems: N-pHLIP in a DOPC and in a POPC bilayer. Both membranes consist of 512 lipids and 6000 waters. The simulation time is 1 μs for each system and the simulation parameters are as described in the Methodology section. The peptide is initially placed in the water phase and quickly adopts an interfacial position. For both systems N-pHLIP remains in a horizontal interfacial orientation during the whole length of the simulations. We also perform MD simulations with N-pHLIP initially half-inserted in a POPC lipid bilayer. Snapshots of the simulations are shown in Figure 17. In the figure, water is represented by blue beads, lipid heads by purple beads and the backbone of the peptide is colored orange. The hydrophilic side chains of the peptide are shown in green. Initially half-inserted in the bilayer, the peptide creates a pertubation to the bilayer after a few nanoseconds and finally adopts a transmembrane orientation in which it stays until the end of the simulation (350 ns).

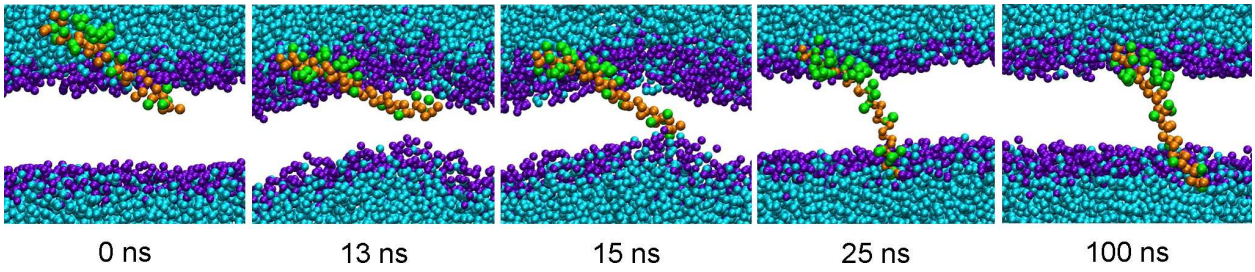


Figure 17: Snapshots from the N-pHLIP simulation. The peptide is initially half-inserted in the lipid bilayer (*left*). As the simulation evolves the peptide perturbs the bilayer (*center*), adopts a transmembrane orientation and stays there until the end of the simulation (*right*). The backbone beads are colored orange, water is colored blue, the phospholipid heads are represented by purple beads and the hydrophilic side chains beads of N-pHLIP are represented by green beads. The hydrophilic residues are not presented for clarity. The beads are not to scale.

TMX-1 is another interesting example of transmembrane helix. It has been synthesized to test to what extent it is possible to design helices that insert spontaneously in a lipid bilayer.[44] TMX-1 (WNALAAVAAALAAVAAALAAVAASKSKSKSK), has a 21-residue non-polar core, N- and C-caps, and a highly polar C-terminus. It has been shown to adopt an α-helical secondary structure in the lipid environment and insert spontaneously across the lipid membranes with 50% probability. Here, we carry out MD simulations with TMX-1 in a DOPC lipid bilayer (128 lipids, 1500 waters, 4 Cl-). The total simulation time is 2.3 μs, and the peptide maintains an interfacial orientation during the whole simulation. In Figure 18, we present a characteristic position of TMX-1 during our simulation: the polar C-terminus of the helix remains in contact with water and the non-polar central part is hidden in the lipid heads area. The N-terminus also appears to prefer a position close to the bilayer/water interface probably due to its tryptophan residue.

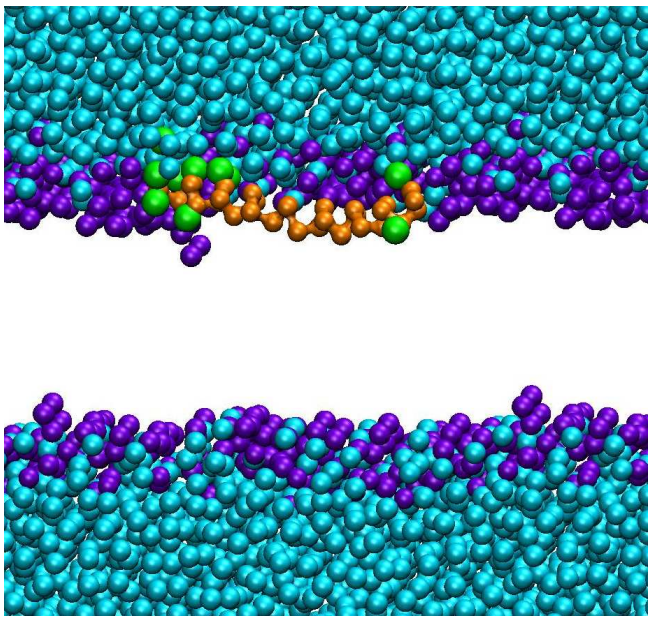
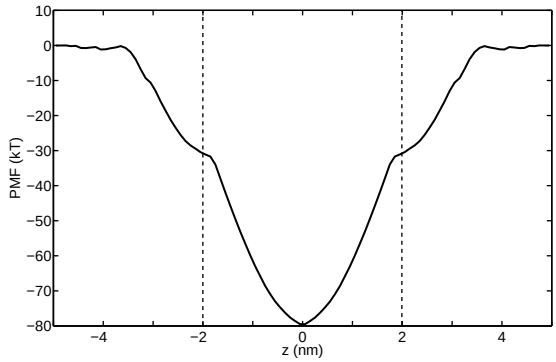


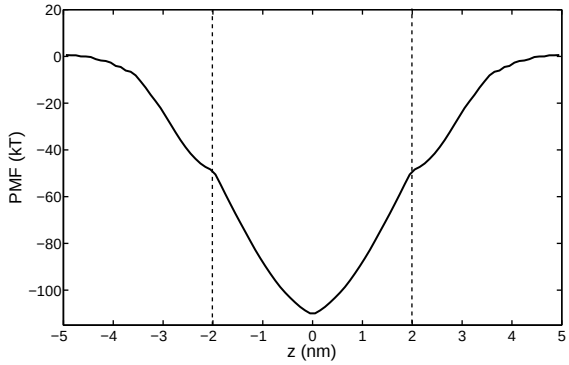
Figure 18: Snapshot from the TMX-1 simulation. The peptide adopts an interfacial orientation with a characteristic orientation where its polar C-terminus remains in contact with water and its non-polar central part is hidden in the lipid heads area. The backbone beads are colored orange, water is colored blue, the phospholipid heads are represented by purple beads and the hydrophilic side chains beads are represented by green beads. The hydrophilic residues are not presented for clarity. The beads are not to scale.

We perform umbrella sampling simulations for both N-pHLIP and TMX-1 peptides. The lipid bilayers used for these simulations are a POPC lipid bilayer for N-pHLIP and a DOPC lipid bilayer

for TMX-1. Both membranes consist of 512 lipids and 12000 water molecules as well as chloride ions to maintain the system neutrality. In Figure 19, we present the resulting PMFs. Again, the generated curves seem to exhibit shape specific for this particular class of peptides. Both peptides show very strong preference for the transmembrane position (with minima at the center of the bilayer at -80kTs and almost -120kTs for pHLIP and TMX-1 respectively) and both PMFs are quite similar, indicating that indeed N-pHLIP and TMX-1 must belong to the same class of peptides. Moreover, these PMFs indicate that one would expect to observe a spontaneous, seamless insertion of those peptides in a lipid bilayer in a MD simulation. This however is not the case. Although peptides, whose orientation is steered into a transmembrane one, remain quite stable in this orientation, we never observe a spontaneous insertion of N-pHLIP or TMX-1. This discrepancy in our observation will be addressed in more detail in the Discussion section.



(a) pHLIP



(b) TMX-1

Figure 19: Potential of mean force for the transfer of N-pHLIP and TMX-1 transmembrane helices across POPC and DOPC lipid bilayers respectively.

Discussion

The results reported in this study suggest that the MARTINI coarse-grained model is able to describe several classes of interactions between α-helical peptides and lipid bilayers. The self-assembly of barrel-stave and toroidal pores, the dynamics of these structures within the bilayer, and the behavior of interfacial non-spanning peptides are within the scope of MARTINI. Furthermore, the model is able to provide new important insights regarding the details of peptide self-assembly in the vicinity of a lipid bilayer. Specifically, using this approach we establish an interesting relation between the barrel-stave and toroidal mechanisms of pore formation. LS3 peptide, being sufficiently long to span the membrane and adopt a transmembrane orientation, is able to form well-defined barrel-stave pores within the bilayer. The number of helices constituting the pore as well as its effective diameter predicted from our simulations is in agreement with previous experimental studies.[31,58] An estimate of the lateral self-diffusion coefficient of the pore gives a value of $D_{lat} = 0.45\ \mu m^2/s$, which is very reasonable, especially taking into account reported values for membrane proteins, such as bacteriorhodopsin.[64] Remarkably, the simulations of a shorter version of this peptide, $(LSSLLSL)_2$, show that this peptide has a much lower propensity for poration and, when a complex does form, it has a toroidal structure. Moreover, the toroidal pores formed by $(LSSLLSL)_2$ are less stable than the complexes formed by LS3. This link between the length of the peptides and their ability to form pores of certain shape and stability is important for the design of pore-forming peptides with specific properties. It will be also of a great interest to establish how these mechanisms depend on other parameters of the system such as, for example, composition of the lipid bilayer. To establish this link a series of extensive simulations was required (totaling to more than 200 μs of simulation time). Clearly, this task could not easily be achieved using a fully atomistic approach.

We perform potential of mean force (PMF) calculations for each peptide considered in this article. In Figure 20, we present a summary of the different PMFs. The PMF provides an estimate of the free energy profile as the peptide crosses the bilayer, and can be used directly to calculate the partition of the peptide between the water phase and the lipid phase. We show that each class

of peptide-bilayer interaction has a very distinct form of PMF (Figure 20). For example, the PMF for LS3 features a deep energy minimum at the center of the bilayer with two additional minima at the interfaces of the bilayer. This PMF is very similar to that observed by Bond and co-workers for WALP23.[66] The PMF for $(LSSLLSL)_2$ is dramatically different with the energy minima at the bilayer interfaces and the core of the bilayer being an unfavorable location of the peptide. In the light of this PMF, the strong propensity of $(LSSLLSL)_2$ for the interfacial orientation in MD simulations is not surprising. The features of the PMF profile for $(LSSLLSL)_2$ are even further magnified for non-spanning LAP20 peptides, that shows even stronger preference for the interfacial location. A very distinct PMF is observed for the fusion SIV peptide, with the energy minimum corresponding to the lipid interface and the energetic penalty for SIV's location in the center of the bilayer of about 10 kT. Finally, the pHLIP and TMX-1 peptides exhibit barrier-less PMFs with -80 kT and -120 kT minima respectively at the center of the bilayer, signifying a very strong preference for the transmembrane orientation. Although this result agrees with our expectations and experimental observations, it contradicts MD studies, where no spontaneous insertion of the peptides in the bilayer is observed. Let us briefly explore possible sources of the discrepancy.

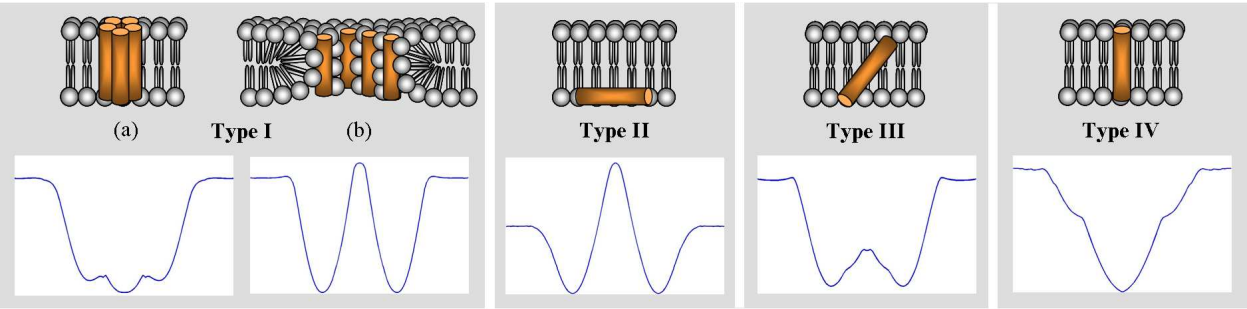


Figure 20: Map of PMF types for the different classes of α-helical peptides.

First, we would like to ensure that the observed PMF is not an artifact resulting from the technical limitations of the applied methods. Specifically, the system features large enough water phase to eliminate possible periodic boundary condition effects. In the original setup, we use the Berendsen thermostat to control the temperature of the system and this method has been criticized for not being able to provide correct distribution of velocities. For one of the smaller peptides,

we repeat the simulations with the Nosé-Hoover thermostat and observe marginal differences. We apply a simple cut-off and shift procedure to electrostatic interactions and this could possibly be viewed as inaccurate. Thus, we recalculate the PMF for one of the smaller peptides, using Particle Mesh Ewald method (PME) to treat electrostatics interactions and observe rather small effects on the final results (see Supporting Information). Moreover, the length of sampling for each window in the Umbrella protocol is important. Most likely the energy profiles reported for large pHLIP and TMX-1 peptides are reflections of a few preferential orientations of the peptides, rather than a result of properly sampled configurational space. Whether this is the case is currently under investigation. However, striking similarity between PMFs for pHLIP and TMX-1 suggests that at least the shape of these energy profiles is characteristic for this particular class.

With these reservations regarding the last two PMFs, we nevertheless believe that the PMF analysis can be used to complement the original classification of peptide-membrane interactions shown in Figure 1 as well as to reveal new types of behavior. Most importantly, it is evident that the PMF analysis is an indispensable tool to elucidate and explain intimate links between different classes of peptides and an exhaustive study of these links would be impossible in atomistic simulations.

Despite several clear successes of MARTINI in application to the aforementioned classes of peptide-membrane interactions, direct extension of this approach (or any other CG approach) to new classes of peptides should be approached with caution. Failure of MARTINI and many other variations of this forcefield to describe the behavior of fusion peptides is just one example of potential difficulties. It seems application of a CG approach to a new class of peptide-membrane interaction should involve careful validation (and if necessary re-calibration) of the approach against known experimental observations for several reference systems within the class. Once the applicability of the CG model is established, the CG approach can indeed provide a number of valuable insights on the behavior of the system as a function of various system parameters.

Other issues with MARTINI-like approaches have been discussed elsewhere[8,66] and we should merely reiterate some of these concerns in the light of new results. MARTINI forcefield implies

that the secondary structure of the peptides is known and does not change in the process of interest. For example, in this study all the peptides are assumed to be α-helical. For many classes of peptides this description is not adequate, as complex conformational evolution of a peptide is often an integral part of the peptide-membrane interaction mechanism. To a significant extent this is also a challenge for fully atomistic models, where accurate description of peptide conformational dynamics in various environments still remains a subject of intense development. Furthermore, the CG treatment of water where several disjoint degrees of freedom are fused together in a single Lennard-Jones site also presents a well recognized issue.[66,72] 'Structureless' and chargeless water is unable to describe water ordering in the vicinity of hydrophobic surfaces or orientation of water molecules in electric fields. This leaves a number of important peptide-membrane processes beyond the scope of MARTINI. For example, the higher propensity of LS3 to form barrel-stave pores in the presence of a weak transmembrane potential most likely will not be adequately described by MARTINI. Translocation of cationic peptides, such as many examples from the family of cell-penetrating peptides, also requires accurate description of water properties. Thus, membrane bending and micropinocytosis, observed for cationic TAT and Penetratin peptides in recent atomistic studies presents a challenge for MARTINI. However, more sophisticated models of CG water started to emerge recently and may provide a solution to at least some of these issues.[73,74]

Even with these limitations, MARTINI-like CG approaches provide a computationally powerful tool to identify and elucidate mechanisms of peptide-membrane interactions and self-assembly processes. The potential structures identified in the CG simulations can then serve as a starting point and be refined in more detailed atomistic studies. Several studies have already addressed a possibility of consistent and systematic multiscale models, where different levels of detail are employed as required at specific stages of the simulation, and this will undoubtedly be one of the most promising areas of development in the future.[75,76]

Acknowledgement

This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF). (http://www.ecdf.ed.ac.uk/). The ECDF is partially supported by the eDIKT initiative (http://www.edikt.org.uk). The authors would also like to thank Ms Naomi Cessford for her diligent proof-reading.

Supporting Information Available

Free energy profiles for the translocation of SIV fusion peptide across a DOPC bilayer calculated by using different approaches for the treatment of electrostatics and temperature coupling. In particular, we perform an umbrella sampling simulation with the Nosé-Hoover thermostat and one with the Particle-mesh Ewald (PME) for the treatment of electrostatic interactions. The information consists of details of the simulations and a comparative figure with the calculated PMFs. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Leontiadou, H.; Mark, A. E.; Marrink, S. J. *J. Am. Chem. Soc.* **2006**, *128*, 12156–12161.

(2) Sengupta, D.; Leontiadou, H.; Mark, A. E.; Marrink, S.-J. *Biochim. Biophys. Acta-Biomembr.* **2008**, *1778*, 2308–2317.

(3) Herce, H. D.; Garcia, A. E. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 20805–20810.

(4) Yesylevskyy, S.; Marrink, S.-J.; Mark, A. E. *Biophys. J.* **2009**, *97*, 40–49.

(5) Izvekov, S.; Voth, G. *J. Chem. Theory Comput.* **2006**, *2*, 637–648.

(6) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750–760.

(7) Venturoli, M.; Sperotto, M. M.; Kranenburg, M.; Smit, B. *Phys. Rep* **2006**, *437*, 1–54.

(8) Marrink, S. J.; de Vries, A. H.; Tieleman, D. P. *Biochim. Biophys. Acta-Biomembr.* **2009**, *1788*, 149–168.

(9) Murtola, T.; Bunker, A.; Vattulainen, I.; Deserno, M.; Karttunen, M. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1869–1892.

(10) Marrink, S. J.; Mark, A. E. *Biophys. J.* **2004**, *84*, 3894–3900.

(11) Marrink, S. J.; Risselada, J.; Mark, A. E. *Chem. Phys. Lipids* **2005**, *135*, 223–244.

(12) Knecht, V.; Marrink, S. J. *Biophys. J.* **2007**, *92*, 4254–4261.

(13) Bond, P. J.; Sansom, M. S. P. *J. Am. Chem. Soc.* **2006**, *128*, 2697–2704.

(14) Bond, P. J.; Holyoake, J.; Ivetac, A.; Khalid, S.; Sansom, M. S. P. *J. Struct. Biol.* **2007**, *157*, 593–605.

(15) Khalid, S.; Bond, P. J.; Holyoake, J.; Hawtin, R. W.; Sansom, M. S. P. *J. R. Soc. Interface* **2008**, *5*, S241–S250.

(16) Balali-Mood, K.; Bond, P. J.; Sansom, M. S. P. *Biochemistry* **2009**, *48*, 2135–2145.

(17) Cox, K.; Sansom, M. S. P. *Mol. Membr. Biol.* **2009**, *26*, 205–214.

(18) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. *J. Chem. Theory Comput.* **2008**, *4*, 819–834.

(19) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.

(20) Risselada, H. J.; Marrink, S. J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2056–2067.

(21) Khalfa, A.; Treptow, W.; Maigret, B.; Tarek, M. *Chem. Phys.* **2009**, *358*, 161–170.

(22) Gkeka, P.; Sarkisov, L. *J. Phys. Chem. B* **2009**, *113*, 6–8.

(23) Brasseur, R. *J. Biol. Chem.* **1991**, *266*, 16120–16127.

(24) Horth, M.; Lambrecht, B.; Chuah Lay Kim, M.; Bex, F.; Thiriart, C.; Ruysschaert, J.-M.; Burny, A.; Brasseur, R. *The EMBO Journal* **1991**, *10*, 2747–2755.

(25) Schanck, A.; Peuvot, J.; Brasseur, R. *Biochemical and Biophysical Research Communications* **1998**, *250*, 12–14.

(26) Oblatt-Montal, M.; Buhler, L.; Iwamoto, T.; Tomich, J.; Montal, M. *J. Biol. Chem.* **1993**, *268*, 14601–14607.

(27) Reddy, G.; Iwamoto, T.; Tomich, J.; Montal, M. *J. Biol. Chem.* **1993**, *268*, 14608–14615.

(28) Becker, C.; Oblatt-Montal, M.; Kochendoerfer, G.; Montal, M. *J. Biol. Chem.* **2004**, *279*, 17483–17489.

(29) Mutter, M.; Tuchscerer, G.; Miller, C.; Altmann, K.; Carey, R.; Wyss, D.; Labhartdt, A.; Rivier, J. *J. Am. Chem. Soc.* **1992**, *114*, 1463–1470.

(30) Grove, A.; Mutter, M.; Rivier, J.; Montal, M. *J. Am. Chem. Soc.* **1993**, *115*, 5919–5924.

(31) Lear, J. D.; Wasserman, Z. R.; Degrado, W. F. *Science* **1988**, *240*, 1177–1181.

(32) Epand, R. F.; Martin, I.; Ruysschaert, J.-M.; Epand, R. *Biochem. Biophys. Res. Commun.* **1994**, *205*, 1938–1943.

(33) Colotto, A.; Martin, J.; Ruysschaert, J.-M.; Sen, A.; Epand, R. M. *Biochemistry* **1996**, *35*, 980–989.

(34) Kamath, S.; Wong, C. *Biophys. J.* **2002**, *83*, 135–143.

(35) Bosch, M.; Fargnoli, K.; Picciafuoco, S.; Giombini, F.; Wong-Staal, F.; Franchini, G. *Science* **1989**, *244*, 694–697.

(36) Martin, I.; Defrise-Quertain, F.; Mandieau, V.; Nielsen, N. M. *Biochemistry* **1991**, *39*, 6581–6585.

(37) Brasseur, R.; Lorge, P.; Goormaghtigh, E.; Ruysschaert, J.-M.; Espion, D.; Burny, A.

(38) Gallaher, W. R. *Cell* **1987**, *50*, 327–328.

(39) Smith, S.; Song, D.; Shekar, S.; Groesbeek, M.; Ziliox, M.; Aimoto, S. *Biochemistry* **2001**, *40*, 6553–6558.

(40) Reshetnyak, Y.; Andreev, O.; Lehnert, U.; Engelman, D. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 6460–6465.

(41) Andreev, O. A.; Dupuy, A. D.; Segala, M.; Sandugu, S.; Serra, D. A.; Chichester, C. O.; Engelman, D. M.; Reshetnyak, Y. K. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 7893–7898.

(42) Reshetnyak, Y. K.; Andreev, O. A.; Segala, M.; Markin, V. S.; Engelman, D. M. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 15340–15345.

(43) Zoonens, M.; Reshetnyak, Y. K.; Engelman, D. M. *Biophys. J.* **2008**, *95*, 225–235.

(44) Wimley, W.; White, S. *Biochemistry* **2000**, *39*, 4432–4442.

(45) Berendsen, H. J. C.; van der Spöel, D.; Vandrunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.

(46) `http://www.hyper.com`.

(47) `http://md.chem.rug.nl/~marrink/MARTINI/Parameters.html`.

(48) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(49) Berger, O.; Edholm, O.; Jahnig, F. *Biophys. J.* **1997**, *72*, 2002–2013.

(50) Siu, S. W. I.; Vacha, R.; Jungwirth, P.; Bockmann, R. A. *Journal of Chemical Physics* **2008**, *128*, year.

(51) `http://www.bioinf.uni-sb.de/RB`.

(52) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular forces*; Pullman, B., Ed.; D. Reidel; Dordrecht, Holland, 1981; pp 331–342.

(53) Bradshaw, J. P.; Darkes, M. J. M.; Haroun, T. A.; Katsaras, J.; Epand, R. M. *Biochemistry* **2000**, *39*, 6581–6585.

(54) de Vries, A. H.; Mark, A. E.; Marrink, S. J. *J. Phys. Chem. B* **2004**, *108*, 2454–2463.

(55) Torrie, G.; Valleau, J. *J. Comput. Phys.* **1977**, *23*, 187–199.

(56) Kumar, S.; Bouzida, D.; Swendsen, R.; Kollman, P.; Rosenberg, J. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(57) Pownall, H. J.; Hu A. amd Gotto, A. M. J.; Albers, J. J.; Sparrow, J. T. *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 3154–3158.

(58) Åkerfeldt, K. S.; Lear, J. D.; Wasserman, Z. R.; Chung, L. A.; Degrado, W. F. *Acc. Chem. Res.* **1993**, *26*, 191–197.

(59) Kienker, P. K.; Degrado, W. F.; Lear, J. D. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4859–4863.

(60) Lear, J. D.; Schneider, J. P.; Kienker, P. K.; DeGrado, W. F. *J. Am. Chem. Soc.* **1997**, *119*, 3212–3217.

(61) Dieckmann, G. R.; Lear, J. D.; Zhong, Q. F.; Klein, M. L.; DeGrado, W. F.; Sharp, K. A. *Biophys. J.* **1999**, *76*, A200–A200.

(62) Chung, L.; Lear, J.; Degrado, W. *Biochemistry* **1992**, *31*, 6608–6616.

(63) Espenel, C.; Margeat, E.; Dosset, P.; Arduise, C.; Le Grimellec, C.; Royer, C. A.; Boucheix, C.; Rubinstein, E.; Milhiet, P.-E. *J. Cell Biol.* **2008**, *182*, 765–776.

(64) Gambin, Y.; Lopez-Esparza, R.; Reffay, M.; Sierecki, E.; Gov, N.; Genest, M.; Hodges, R.; Urbach, W. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 2098–2102.

(65) Randa, H.; Forrest, L.; Voth, G.; Sansom, M. *Biophys. J.* **1999**, *77*, 2400–2410.

(66) Bond, P. J.; Wee, C. L.; Sansom, M. S. P. *Biochemistry* **2008**, *47*, 11321–11331.

(67) Efremov, R.; Nolde, D.; Konshina, A.; Syrtcev, N.; Arseniev, A. *Curr. Med. Chem.* **2004**, *11*, 2421–2442.

(68) Bradshaw, J. P.; Darkes, T. A.; Katsaras, J.; Epand, R. M. *Physica B* **2000**, *276-278*, 495–498.

(69) Harroun, T. A.; Balali-Mood, K.; Gourlay, I.; Bradshaw, J. P. *Biochim. Biophys. Acta-Biomembr.* **2003**, *1617*, 62–68.

(70) Martin, I.; Dubois, M.-C.; Defrise-Quertain, F.; Saermark, T.; Burny, A.; Brasseur, R.; Ruysschaert, J.-M. *J. Virol.* **1994**, *68*, 1139–1148.

(71) Lins, L.; Charloteaux, B.; Thomas, A.; Brasseur, R. *Proteins* **2001**, *44*, 435–447.

(72) Bock, H.; Gubbins, K. E.; Klapp, S. H. L. *Phys. Rev. Lett.* **2007**, *98*, year.

(73) Iuchi, S.; Izvekov, S.; Voth, G. A. *J. Chem. Phys.* **2007**, *126*, year.

(74) Orsi, M.; Haubertin, D. Y.; Sanderson, W. E.; Essex, J. W. *J. Phys. Chem. B* **2008**, *112*, 802–815.

(75) Chu, J. W.; Ayton, G. S.; Izvekov, S.; Voth, G. A. *Mol. Phys.* **2007**, *105*, 167–175.

(76) Carpenter, T.; Bond, P. J.; Khalid, S.; Sansom, M. S. P. *Biophys. J.* **2008**, *95*, 3790–3801.