School of Pharmacy and Biomedical Science

Assessment of a Novel Enhanced Sampling Method for Molecular Simulations of Peptide-membrane Interactions.

John Douglas Tanner
0000-0003-3914-3116

This thesis is presented for the Degree of
Master of Research (Biomedical Science)
of
Curtin University

May 2020

Declaration
To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.
This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature:
Date: 11/05/2020

Acknowledgements
This research is supported by an Australian Government Research Training Program (RTP) Scholarship and is also supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia.

Abstract

Anti-microbial and anti-cancer peptides are of interest for their therapeutic potential. Many of these peptides predominately act by permeabilising cell membranes but the detailed molecular mechanism is still not fully understood, hindering the rational design of clinically useful peptides. Molecular dynamics (MD) simulations can offer a molecular-level insight into the mechanisms of action of these peptides, including the prediction of their free energy of binding (ΔG_b) to membranes. However, to accurately calculate ΔG_b, MD simulations need to sample both low- and high-energy states. In the case of peptide-membrane systems, this involves sampling an extensive range of configurations of the peptide on the membrane surface, a task difficult to achieve using conventional MD. To address this, we use a modified version of the enhanced sampling method Replica Exchange with Solute Scaling (REST2), REST3, where the electrostatic and van der Waals interactions of peptide-membrane and peptide-water are selectively scaled to enhance configurational sampling of the peptide. REST3 was developed in the Curtin Biomolecular modelling group lab group in conjunction with the Curtin Institute for computation. To validate REST3, the peptide Gomesin was used as a test system. Gomesin is an 18-residue peptide that shows potent cytotoxic activity against clinically relevant bacteria and fungi, as well as human and murine cancers, and the peptide acts by binding to cell membranes. Gomesin has a rigid β-hairpin structure, reinforced with two cross-strand disulfide bonds, and consequently exhibits limited conformational changes upon membrane binding. To calculate ΔG_b for the interaction of Gomesin to a model cell membrane

consisting of a phospholipid bilayer, REST3 was combined with umbrella sampling. Use of REST3 was found to increase configurational sampling of the peptide; however, the predicted value of ΔG_b is inaccurate compared the value predicted by only using umbrella sampling. Detailed analysis suggests that REST3 causes undesirable changes in the balance of interactions between different system components. The most significant effect is the anomalous formation of a layer of water at the peptide-membrane interface, interfering with the formation of peptide-membrane interactions. An extensive range of strategies were investigated to try to mitigate the problem.

Keywords: peptide-membrane interactions, molecular dynamics simulations, anticancer peptides, antimicrobial peptides, structure-based drug design, enhanced sampling

Contents

1.0 Introduction **5**

2.0 Methods **15**

2.1 Simulation systems and equilibration 15

2.2 Overview of simulations 16

2.3 Conventional umbrella sampling simulations 18

2.4 Simulations to determine the optimal scaling factors 19

2.5 REST3-umbrella sampling simulations 19

2.6 Assessment of configurational sampling 20

2.7 Hydrogen bond analysis 21

2.8 Calculation of solvation index 21

3.0 Results and Discussion **23**

3.1 Determination of windows that require REST3 23

3.2 Optimisation of scaling factors for REST3 26

3.3 Effect of scaling on the interactions in the system 30

3.4 Effect of maximum scaling factor on replica exchange simulations 33

3.5 Potential of mean force calculations 39

3.6 Determination of the desolvation timescale of Gomesin 42

3.7 Peptide-peptide interaction scaling 45

3.8 Membrane-water interaction scaling 46

3.9 Effect of unequal peptide-membrane and peptide-water scaling 47

4.0 Conclusions **53**

5.0 References **57**

6.0 Supporting Information **62**

1.0 Introduction

Anti-microbial and anti-cancer peptides (AMPs and ACPs) are a promising avenue for new therapeutic drugs to combat multi-drug resistance in both bacteria and cancer cells.(1, 2) Both AMPs and ACPs are typically cationic and amphiphilic.(3) The primary mechanism of their cytotoxic activity is through the binding to and permeabilisation of the plasma and mitochondrial membranes of the target cell.(4) Though promising, a lack of understanding of the molecular factors controlling membrane binding and permeabilisation hinders a rational design approach for produce therapeutically useful peptides. To better apply rational design approaches to membrane disruptive peptides, a molecular level understanding of the peptide-membrane interactions is required.(5) Specifically, it is necessary to know how peptides bind, enter and disrupt membranes. This in turn requires an understanding of the preferred orientation(s) of peptides on the membrane and identifying the amino acid residues that govern membrane binding. Molecular dynamics (MD) simulations can provide this information by studying the structure and dynamics of biomolecular systems at the molecular level. MD simulations can also predict thermodynamic properties such as the free energy of binding (ΔG_b) to membranes.(6, 7) However, the accurate and reliable calculation of ΔG_b is one of the most challenging tasks in biomolecular simulations and requires validation by comparison to experimental data from matching peptide-membrane systems.(8)

One of the most commonly used methods to obtain ΔG_b from MD simulations is umbrella sampling (US).(9) In this method a series of independent simulations (a.k.a windows) are performed to calculate the change in free energy along a path that connects two states, referred to as the reaction coordinate. For peptide-membrane systems, these two states are usually defined as the peptide in solution and the peptide in its membrane-bound state, and the reaction coordinate is the distance between the peptide and membrane centres of mass (Fig. 1A). For each window, the peptide is restrained at a specific distance to the membrane and data from all windows is subsequently combined to reconstruct the potential of mean force (PMF), which describes the relative free energy as a function of the reaction coordinate. From the PMF, ΔG_b can be estimated as the difference between the free energy in the bulk solution and the free energy at the interface. For ΔG_b to be accurate, each window needs to sample a sufficiently large number of representative low- and high-energy

configurations of the system. In practice, this is difficult to achieve in peptide-membrane simulations because at short peptide-membrane distance, the rotational and translational motion of the peptide is greatly hindered. Consequently, US simulations usually will not sufficiently sample the configurational space within accessible time scales.(10)

This problem has sought to be addressed through the use of a number of enhanced sampling techniques such as coarse Grained (CG) models, metadynamics (MetaD), and replica exchange methods.(11) No approach taken so far has yet proven to be suitable for increasing the configurational sampling and providing accurate estimations of ΔG_b for peptide-membrane systems. Each of these methods has advantages, but also disadvantages that make their application to peptide-membrane systems unsuitable.

MetaD is a method (or more accurately, a family of methods) in which previously visited states are biased against using a history-dependent bias.(12) These states are defined as a function of one or more 'collective variables' (CVs), which can be any differentiable function of atomic Cartesian coordinates. Early studies of peptide-membrane systems using classical MetaD generally only used a single CV defined by the peptide-membrane COM distance. This, unfortunately, can produce inaccurate PMFs, as the relaxation time of membrane deformations caused by the movement of the peptide is much longer than the rate at which the peptide-membrane COM distance changes.(13) These deformations of the membrane caused the peptide were found to take hundreds of nanoseconds to re-equilibrate. This issue makes the accurate estimation of the ΔG_b of AMP/ACP binding to membranes using MetaD very challenging, as these peptides typically strongly interaction and deform membranes.(14) However, Cao et al. recently used bias-exchange MetaD to study five-residue long peptides that spontaneously translocate membranes. (15, 16) To bypass the problem of not sampling a wider range of configurations because the peptide motion is dominated by strong interactions with the membrane, a total of six collective variables were used. In addition to the peptide-membrane COM distance, CVs for the number of contacts between each of the five peptide residues and the membrane were also included. This allowed the peptides to lose and regain interactions with the membranes on a feasible timescale. The calculated free energies of membrane crossing were highly correlated with the experimentally known translocation rates of these peptides. While the free energy calculations were successful, this approach is not

suitable for larger peptides as MetaD simulations with large numbers of CVs are too computationally expensive and tend to be inefficient.

CG models can increase the sampling of membrane surface interactions by reducing the degrees of freedom in the system. In this approach several atoms are represented by a single particle (called CG 'bead'). This reduces the number of particles in the system and allows a larger timestep to be used, which in turn means that simulations of tens of microseconds are feasible.(17) The longer simulation times allow the peptide(s) sufficient time to reorient and sample new configurations.(18) However, CG particles do not provide the same resolution of interactions. For example, hydrogen bonding between peptide side chains and lipid headgroups cannot explicitly be described by CG particles. As such, atomistic-level descriptions of mechanisms of interaction can be limited in these simulations and estimations of the free energies of the underlying processes may be less accurate. CG models with implicit solvents will also not capture in detail the important interactions between the peptide and immediately neighbouring water molecules.(19) Secondary structure changes are also unable to be modelled with most CG models as the peptide structure is usually constrained within an elastic network model.(20) In the case of peptides with a rigid beta-hairpin structure that is the same in an aqueous or lipid environment, this approach is permissible. However, for α-helical AMPs such as melittin, capturing the transition from a random coil in solution to a α-helix on the membrane is important for an accurate calculation of the free energy of interaction.(21-23)

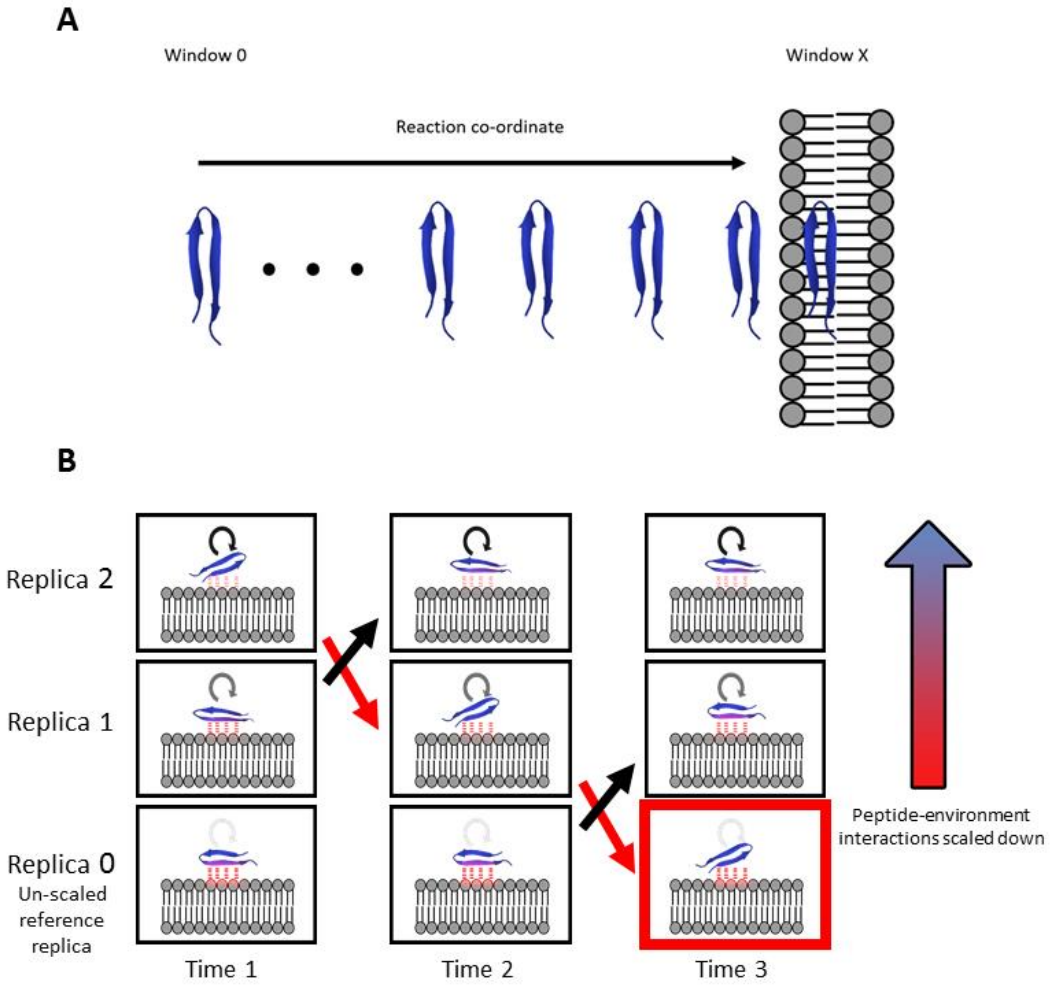


Figure 1: Diagram of enhanced sampling techniques in peptide-membrane simulations. A: In umbrella sampling the peptide is studied in a series of independent simulations, known as windows. In each window the peptide is at a different distance to the membrane along the reaction coordinate, which follows the peptide in solution to its membrane-bound state. B: In REST3 each window contains several replicas of the system. Each replica has a different Hamiltonian. The Hamiltonian in the reference replica is unchanged. As the replica 'ladder' is climbed, the Hamiltonian is modified to a larger degree. In REST3, the Hamiltonians have been modified so that the interactions of the peptide with its environment are reduced in magnitude (i.e. they are scaled down), enhancing the sampling of peptide configurations. At regular intervals, exchanges of structures are made between the replicas. This allows the new configurations of the peptide found in the higher replicas to propagate down into the reference replica

The final class of enhanced sampling methods discussed here are known as replica exchange methods. The approach used in this study belongs to this class of methods. Selective replica-exchange with solute tempering (REST3) is a method developed in-house that is an extension of the original replica-exchange molecular dynamics.(24) In replica-exchange methods, a simulation contains several replicas of the system that are

run concurrently. In each replica, specific interactions in the system are tempered, with the reference replica experiencing no tempering, and the highest replica experiencing the most tempering (Fig. 1B). At regular intervals, exchanges of the configurations (X) between neighbouring replicas (*m* and *n*) are attempted as described below:

$$X_m, E_m(X_m), T_m \rightarrow X_n, E_m(X_n), T_m$$
$$X_n, E_n(X_n), T_n \rightarrow X_m, E_n(X_m), T_n \quad \text{(Eq. 1)}$$

where the Hamiltonian, E, and temperature, T, of the replica remain the same after exchanges. An exchange will occur with an acceptance probability determined by applying the detailed balance condition to the transition above, such that

$$P_m(X_m)\, P_n(X_n)\, P(X_m \rightarrow X_n) = P_m(X_n)\, P_n(X_m)\, P(X_n \rightarrow X_m) \quad \text{(Eq. 2)}$$

and given that the probability of a configuration is

$$P_m(X_m) = e^{-\beta_m E_m(X_m)} \quad \text{(Eq. 3)}$$

where β = 1 / kT and k is the Boltzmann constant. The ratio of transition probabilities is

$$\frac{P(X_m \rightarrow X_n)}{P(X_n \rightarrow X_m)} = e^{-\Delta_{nm}} \quad \text{(Eq. 4)}$$

where:

$$\Delta_{nm} = -\beta_m[E_m(X_n) - E_m(X_m)] - \beta_n[E_n(X_m) - E_n(X_n)] \quad \text{(Eq. 5)}$$

The Metropolis criteria is then used to calculate the probability of the exchange being accepted, P(a):

$$P(a) = \begin{cases} 1 & if\ \Delta_{nm} \leq 0 \\ \exp(\Delta_{nm}) & if\ \Delta_{nm} > 0 \end{cases} \quad \text{(Eq. 6)}$$

In the original replica exchange MD method (T-REMD), replicas are simulated at increasing temperatures, as the larger kinetic energy in the higher temperature replicas enables the crossing of energy barriers and thus access to configurations that would not be accessible in lower replicas.(24) The exchange between replicas then allows these high-energy configurations to 'trickle down' into the ground replica, increasing the set of configurations sampled compared to conventional MD.(24) However, this approach is inefficient for systems where the solute molecule is surrounded by a large number of solvent molecules. In REST this problem is solved by dividing the Hamiltonian of the system into solute-solute, the solvent-solvent and solute-solvent components.(25) All replicas still have an increased temperature, but the solvent is scaled such that the energy component for the solvent-solvent interactions no longer factors into the acceptance probability. The energy difference between replicas is

consequently smaller and exchanges between replicas are more frequent, greatly increasing the efficiency of the method. In the more recent REST2 method, there is no change in temperature and, instead, only the solute of interest (e.g. a protein) is tempered selectively, leaving the solvent (usually water) unchanged.(26) The Hamiltonian is decomposed as such:

$$E_m(X) = \frac{\beta_m}{\beta_0} E_{pp}(X) + \sqrt{\frac{\beta_m}{\beta_0}} E_{pw}(X) + E_{ww}(X) \qquad \text{(Eq. 7)}$$

where the energy of structure X of the *m*th replica, E_m, is broken up into the solute-solute (protein-protein) terms, *pp*, the solute-solvent (protein-water), *pw*, and solvent-solvent (water-water), *ww*. E_{pp} and E_{pw} are scaled by functions of β_0, which is thermodynamic β at the temperature of the ground replica, and β_m, at T_m. T_m is the so-called 'effective temperature' and this is the parameter that is modified to scale the interactions. Replica *m* is not actually run at this temperature and it remains at T_0, the temperature of the ground replica.

Replica exchange methods were all developed for a single solute of interest and, in particular, for protein folding. Examples of studies using replica exchange methods for peptide-membrane systems are sparse.(27-35) Rather than aiming to study the orientational sampling of the peptide on the membrane surface, replica exchange methods are utilised to enhance the conformational sampling of the structure of peptides in a membrane environment. T-REMD appears to be the most widely utilised replica exchange method in peptide-membrane systems, despite the high temperatures causing structural disturbances in the membrane.(27-32) In some cases, to avoid these temperature-induced membrane disturbances, implicit membrane models have been used, though the accuracy of representing the dynamic membrane environment with a fixed dielectric constant is questionable.(27-30) Compared to T-REMD, in REST/REST2 methods it is possible to only temper the solute of interest, leaving the membrane unaffected. REST has been used to study the conformational sampling of Aβ fragments in the presence of membranes.(34, 35) REST achieves increased conformational sampling by increasing only the temperature of the solute of interest. REST2 has been applied to the conformational sampling of a huntingtin protein fragment that acts as a membrane anchor.(33) As can be seen from Eq. 7, the solute-solute terms are always scaled, and the solute-solvent interactions are always scaled by the square root of this reduction. Reducing the E_{pp} interactions encourages conformational sampling as it affects the structure of the peptide. This is not suitable

for many peptide-membrane systems, and in particular, for AMPs and ACPs, which rely on a well-defined structure for their activity.(14, 34, 35)

The new method tested in this project, REST3, expands REST2 by allowing the selective scaling of any of the components of the Hamiltonian of the system. For a peptide-membrane system, this new method allows to separately scale the peptide-membrane and peptide-water interactions (including, if needed, different levels of scaling) whilst leaving the peptide-peptide, water-water and membrane-membrane interactions unaffected. Eq. 8 demonstrates how the Hamiltonian would be decomposed for a peptide-membrane system:

$$E_m(X) = \lambda_{pp}E_{pp}(X) + \lambda_{pw}E_{pw}(X) + \lambda_{pm}E_{pm}(X) + \lambda_{mw}E_{mw}(X) + \lambda_{mm}E_{mm}(X) + \lambda_{ww}E_{ww}(X) \quad \text{(Eq. 8)}$$

where λ is the 'scaling factor', which can be any real number. Note that the formulation of the Hamiltonian is specific to the components of the system. If additional system components were present, the Hamiltonian would be further partitioned into separate terms for any self- and inter-component interactions in the system. In addition, the van der Waals (vdW) and electrostatic interactions between any two system components can also be scaled separately. In this study, the REST3 formalism was used to selectively scale the peptide-membrane and peptide-water interactions with the aim of more efficiently sampling the configurations of the peptide at the water-membrane interface, whilst leaving intra-peptide interactions unscaled. REST3 has been previously tested with three small terpene molecules and was able to efficiently increase configurational sampling and predict experimental ΔG_b values with improved accuracy. (Martinotti, manuscript in preparation)

To assess the suitability of REST3 for the prediction of ΔG_b in larger, peptide-membrane systems, the binding of the peptide Gomesin (Gm) to a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayer was used as a test case. Gm is an 18-residue peptide that was originally isolated from the immune cells of the tarantula *Acanthoscurria gomesiana.(36)* Gm has the sequence ZCRRLCYKQRCVTYCRGR-NH_2, where Z is a pyroglutamic acid and R-NH_2 is an amidated Arg.(36) The peptide adopts a rigid, β-hairpin-like fold in which anti-parallel β-strands are connected by a 4-residue turn (Fig. 2A).(37) The structure is stabilised by two disulfide bonds, Cys2-Cys15 and Cys6-Cys11, as well as six inter-strand backbone-backbone hydrogen bonds. This rigid structure makes Gm an ideal test case for REST3, as conformational

sampling of the peptide is not desirable. Gm has a +6 charge due to the presence of five Arg and one Lys. The C- and N-termini are neutral due to the amidation and pyroglutamic acid. The presence of a hydrophobic face and charged/polar residues gives the peptide an amphiphilic nature (Fig. 2B).(37) Amphiphilicity, positive charge, and an anti-parallel β-hairpin structure are common characteristics in other peptides such as tachyplesin, polyphemusin and protegrin that act as host defence peptides in the innate immune system, having immuno-modulatory and anti-microbial activity.(38-40)

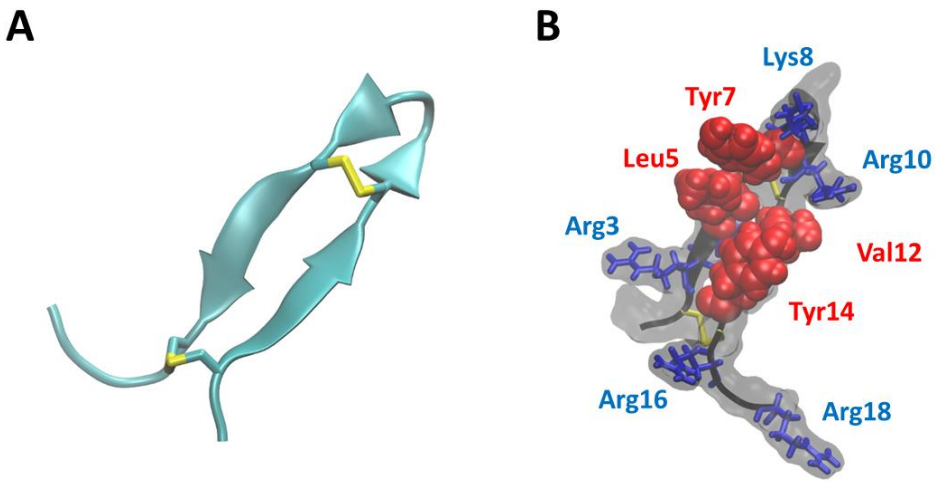


Figure 2: Structure of Gomesin (Gm). A: NMR solution structure of Gm resolved at 298K. Disulfide bonds shown in yellow. B: Representation of Gm showing the hydrophobic (red) and charged (blue) residues. Note that Arg4 is on the other side of the peptide.

Gm is cytotoxic towards bacteria such as *Staphylococcus aureus*, *Salmonella thyphinirium*, and *Pseudomonas aeruginosa*, which are bacteria listed as high priority for the development of new antibiotic treatments by the World Health Organisation.(36, 41) Gm also has anti-cancer activity against breast, colon, leukaemia and melanoma cell lines, and has been shown to reduce tumour growth in mouse models of melanoma.(42-47) Anti-cancer peptides such as Gm can complement existing chemotherapy treatments.(48, 49) The common activity against bacteria and cancer cells can be rationalised by understanding the mechanism of action of Gm and that of structurally-related, cationic, amphiphilic peptides.(49-52) Gm exerts its cytotoxic activity by permeabilising the plasma cell membrane as well as through intracellular effects. The subsequent loss of membrane integrity from the permeabilisation results in cell death. This membrane-disrupting activity of Gm has

been demonstrated in lipid vesicles as well as whole-cell experiments with cancer cells.(43, 45, 53-56). The lytic activity of Gm increases with increasing concentration of negatively charged lipids in membranes, as would be expected from the cationic nature of the peptide.(43) As both bacterial and cancer cells have an increased concentration of negatively charged lipids in the outer leaflets of their plasma membranes compared to healthy human cells, electrostatic attraction is likely to be one of the main factors that controls the selectivity of Gm towards these pathogens.(48, 51, 52, 57, 58)

In general, the ability of Gm to bind to membranes correlates with its anti-cancer and anti-microbial activity.(43) At low concentrations Gm also causes cell death through late apoptosis. However, irrespective of the mechanism, both permeabilisation and apoptosis are preceded by membrane binding.(43, 45, 59) Thus, by understanding the molecular mechanism of membrane binding, it might be possible to design Gm variants with increased cytotoxic activity and better selectivity towards bacterial and cancer cell membranes.

Most experimental studies on the membrane-binding properties of Gm have focused on understanding the role of the β-hairpin structure, the role of charge and hydrophobicity of Gm, or the role of the lipid composition of membranes.(21, 43, 53-55, 60, 61) A large number of peptide variants (also called mutants), in which one or more residues are modified or replaced to alter the physicochemical properties of the peptide, have been prepared and their membrane binding properties were compared to that of native Gm.(43, 53) These studies showed that further increasing the positive charge of the peptide increases its binding to negatively charged membranes, and reducing its hydrophobicity correlates with reduced membrane binding.(43, 53) There are, however, exceptions to these trends, which hint at a more complicated structure-function relationship and that increasing membrane binding is not as simple as maximising the overall charge and/or increasing hydrophobicity. For example, the peptide variants R3A and R10A, where Arg3 or Arg10 is replaced by Ala, have a reduced positive charge compared to native Gm and, despite this, these mutants bind to negatively charged membranes with an increased affinity.(53) In contrast, the mutations R4A and R18A do not affect membrane binding.(43) Similar exceptions exist when mutating residues in the hydrophobic face: the variants Y7A, V12A and Y14A show a complete loss of membrane binding but L5A shows only slightly reduced binding compared to native Gm.(53) These examples illustrate that the

relationship between charge, hydrophobicity and membrane binding is not always straightforward. The effect of these mutations is likely dependent on the position of the peptide residue in relation to the membrane. Hence, characterising the membrane-binding behaviour of Gm using MD simulation will likely shed light into these observations. However, the ΔG_b must first be accurately calculated and validated in order to lend credibility to the mechanistic insights that MD simulation may afford. To do so, the applicability of the REST3 method to the accurate computation of ΔG_b for the Gm-POPC system was tested.

The aim of this project was to apply REST3 to the peptide membrane system of Gomesin with a POPC bilayer and enhance the sampling of configurational states. And from these simulations, obtain the relative binding affinities of Gm to a POPC membrane. With these binding affinities, the simulations may be validated against experimental data. Extending this method to Gm variants and other membrane compositions is a further goal, but not specifically the aim of this project.

This thesis describes in detail the attempts to apply REST3 to the Gm-POPC system. Chapter 3.1 describes the identification of which US windows require the use of the REST3 approach due to poor sampling of peptide-membrane orientations. Chapter 3.2 details the process of how peptide-environment scaling factors were optimised to allow for the maximal sampling of peptide-membrane configurations. In Chapter 3.3 analyses are presented of the unintended side-effects of scaling the peptide-environment interactions on the hydrogen bonding between system components. In Chapter 3.4, the different peptide-environment scaling factors that were incorporated into replica exchange simulations and the effect of this scaling of interactions is assessed. In Chapter 3.5, the scaling factor that was most effective at enhancing the sampling of peptide-membrane configurations was used to calculate a PMF and estimate the ΔG_b. This PMF was found to be inaccurate because of the formation of a water layer between the peptide and membrane surface, effectively preventing the desolvation of the peptide. Chapter 3.6 investigated the amount of time needed to dissipate this layer of water and implements the estimated desolvation time into a replica exchange simulation. The following Chapters 3.7 and 3.8 investigate the alternative scaling approaches of reducing peptide-peptide interactions and membrane-water interactions, respectively, to try to increase the rate of desolvation of the water layer. Finally, in Chapter 3.9 consideration is given to the inclusion of replicas wherein the peptide-environment interactions are scaled unequally.

2.0 Methods

2.1 Simulation systems and equilibration

All simulation systems were built with a pre-equilibrated POPC bilayer composed of 128 lipids (64 per leaflet) and a single peptide, solvated with 12772 water molecules, giving a water:lipid ratio of approximately 100, which corresponds to a fully hydrated lipid bilayer and with a sufficiently large amount of water for the umbrella sampling calculations. A structure of Gm recently determined by NMR spectroscopy at 298 K was used as a starting structure (PDB ID 6MY2).(62) The peptide was modelled with an amidated Arg at the C-terminus and a L-pyroglutamic acid at the N-terminus. L-pyroglutamic acid was modelled by adding the C=O group from a Gln side chain to a Pro residue. Atom types, bonds, angels and dihedrals for L-pyroglutamic acid were based on Gln and Pro residues while the partial charges were calculated using the Automated Topology Builder (ATB).(63) The residue pairs Cys2 – Cys15 and Cys6 – Cys11 were connected by disulfide bonds. Torsion angle restraints were applied to the Cys2 – Cys15 disulfide bond to stabilize the structure of Gm, which the Gromos54a7 and other force fields have difficulty maintaining as discussed in Deplazes et al. (2019). The torsion angles used were -60, -90, 120, -50 and -60 for torsions χ_1, χ_2, χ_{ss}, $\chi_{2'}$ and $\chi_{1'}$, respectively, which were based on values from the NMR structure known at the time.(62) A force constant of 100 kJ mol^{-1} was used to restrain the torsion angles. The peptide charge of +6 was neutralised with 6 Cl^- ions, and a further 52 Na^+ and 52 Cl^- ions were added to bring the ionic strength to 230 mM.

The centre of mass (COM) of the peptide was placed initially in the aqueous environment approximately 7.8 nm away from the COM of the bilayer. The systems were energy minimised using conjugate gradients, followed by a 500 ps simulation in the NVT ensemble and a 500 ps simulation in the NPT ensemble at 298 K and 1 atm of pressure. To allow full equilibration of the membrane in its solvent environment, further NPT simulations were carried out using a harmonic potential (with a force constant of 500 kJ/mol) to keep the peptide at its original position with respect to the membrane. The area per lipid (APL) of the lipid bilayer was used to monitor equilibration. Once a stable APL was reached, the harmonic potential on the peptide was removed and the system was simulated for 150 ns in the NPT ensemble. From this simulation, the starting structures used in the subsequent US and REST3-US simulations were extracted. If the window distance was not present in this simulation, the starting structure was obtained from the nearest window.

All simulations were performed with GROMACS 4.6.7.(64) The GROMOS 54a7 force field was used for the peptide, while POPC was described using the GROMOS-compatible lipid parameters developed by Poger et al.(65, 66) The SPC water model was used.(67) A 2 fs time step was used in all simulations. The system was modelled in a rectangular box with periodic boundary conditions in all directions. The reaction field method was used with a distance cut-off of 1.4 nm to compute electrostatic interactions, as to be consistent with the Gromos force field.(68) The dielectric constant was set to 62 Fm^{-1} beyond the cut-off. A single cut off of 1.4 nm was used for vdW interactions with no long-range dispersion corrections. The Berendsen thermostat and barostat were used during equilibration.(69) The barostat was applied semi-isotropically, with a reference pressure of 1.0 bar (1 atm), a compressibility of 4.6 x 10^{-5} bar^{-1} and a time constant of 1 ps. The thermostat was set to 298 K with a time constant of 0.1 ps. The protein, POPC, and solution (including ions) were coupled separately to the thermostat. The Nose-Hoover thermostat was used during the production runs, with the same temperature of 298 K, which was updated every 0.5 ps.(70) Pressure was maintained at 1.0 bar with the Parrinello-Rahman barostat using a time constant of 2 ps and a compressibility of 4.6 x 10^{-5} bar^{-1}.(71) The SETTLE algorithm was used to constrain the covalent bonds in water, whilst the LINCS algorithm was used for all other covalent bonds.(72, 73) Frames were written every 2 ps.

2.2 Overview of simulations

Table 1: Summary of all simulations conducted in this study. IDs are given to each simulation for in-text cross referencing. Abbreviations: PM – peptide-membrane SF, PW – peptide-water SF, PP – peptide-peptide SF, MW – membrane-water SF. Under scalings, if a particular SF is not mentioned, it should be assumed equal to 1.0. Related simulations have been grouped into 'sets' for the purposes of readability. Each simulation set generally corresponds to simulations discussed in a section of the results.

ID	Simulation set	Description	# of simula-tions	Run time (ns per simulat ion)	Scalings	Peptide-membrane COM Distance (nm)
cUS-2.2 → cUS-6.6	Conventional Umbrella sampling	Conventional MD simulations held at different distance windows for the calculation of a PMF	22	300	NA	2.2 → 6.6 at 0.2 nm increments
SS-0.8 → SS-0.2	Scaled simulations	Simulations of Gm-POPC with different the peptide-environment scaling factors reduced	4	200	PM and PW: 0.8, 0.6, 0.4, 0.2	2.6
R-0.8 → R-0.2	REST3 testing	REST3 simulations where the maximum peptide-environment SF is reduced	4	50	Max SFs of PM and PW: 0.8, 0.6, 0.4, 0.2	2.6
Cr-1 → Cr-10	Conventional replicate simulations	Ten duplicate conventional MD simulations to control for the different amounts of data with respect to REST3 simulations	10	50	NA	2.6
RUS-2.2 → RUS-4.2	REST3 Umbrella sampling	REST3 simulations with a maximum peptide-environment SF of 0.2 for the calculation of PMF	11	50	Max SF of PM and PW: 0.2	2.2 → 4.2 at 0.2 nm increments
CC-1 → CC-5	Continued conventional	Conventional MD simulations started with structures taken at random from RUS-2.6	5	75	NA	2.6
sPP-0.8 → sPP-0.2	Peptide-peptide scaling	Simulations of Gm-POPC with different	3	200	PP: 0.8, 0.6, 0.4,	2.6

		reduced peptide-peptide scaling factors.				
sMW-0.8 → sMW-0.2	Membrane-water scaling	Simulations of Gm-POPC with different reduced membrane-water scaling factors.	4	200	MW 0.8, 0.6, 0.4, 0.2	2.6
sUe-0.8 → sUe-0.2	Unequal SF	Simulations of Gm-POPC where only the peptide-water interactions are reduced	4	200	PM 1.0, 1.0, 1.0, 1.0 PW 0.8, 0.6, 0.4, 0.2	2.6
rME-0.5	Middle-earth REST3	A REST3 simulation where the SF of the replicas are in a 'middle-earth' arrangement	1	50	Unequal SF max SF: PM = 1.0 PW = 0.5 Equal SF max, PM and PW = 0.5	2.6
R-EAF	REST3 EAF test	REST3 simulation with an EAF of 500 ps to test the estimate of the minimum required desolvation time	1	50	Max SFs of PM and PW at 0.2	2.6
SS2.2-0.2 → SS2.2-0.01	Scaled simulations at 2.2 nm	Simulations with peptide-environment SF reduced at a close COM distance of 2.2 nm	5	200	PM and PW SFs of 0.2, 0.1, 0.075, 0.05, 0.01	2.2

Conventional MD simulations typically required 1440 CPU hours per 100 ns of simulation time. Scaled-only simulations typically required 1680 CPU hours per 100 ns of simulation time. REST3 simulations with a max SF of 0.2 required 15,500 CPU hours per 100 ns of simulation time. All simulations were performed on the Magnus system at the Pawsey Supercomputing Centre.

2.3 Conventional US simulations

The peptide-membrane COM-distance along the z-axis (normal to the membrane surface on the X-Y plane) was defined as the reaction coordinate for all US simulations

(see also Fig. 1A). Windows were spaced 0.2 nm apart, with a total of 23 windows ranging from a COM-distance of 2.2 to 6.6 nm. Each window was simulated for 250 ns using the same parameters as for the unbiased MD simulations. The PMF was reconstructed with the weighted histogram analysis (WHAM) method with the GROMACS tool g_wham using only the last 50 ns in the trajectory of each window.(74, 75) Uncertainties were calculated using the bootstrapping algorithm in g_wham, with 200 bootstraps.

2.4 Simulations to determine the optimal scaling factors

A series of scaled simulations with no replica exchange were conducted to determine the optimal scaling (tempering) factors. The system was simulated for 200 ns with scaling factors (SFs) of 0.8, 0.6, 0.4, and 0.2. In all simulations, only the interactions between the peptide and the other components (water, membrane, ions) were scaled, leaving intramolecular peptide interactions, membrane-membrane and membrane-water interactions unscaled. These interactions have been termed the 'peptide-environment' interactions, and if there is only a mention of a single SF, it should be assumed that this is what is being scaled. The same SF was applied to both the vdW and electrostatic interactions. For example, in simulations with a peptide-environment SF of 0.2, both vdW and electrostatic interactions for peptide-membrane, peptide-water and peptide-ion interactions were scaled down using a factor of 0.2 (i.e. their magnitude was 20% of their full value). An example of the scaling matrix used is shown in Table S1. The optimal SF is defined to be the SF that allows for the greatest sampling of configurations, whilst being as close as possible to a value of 1.0, so as to reduce the number of replicas needed in a full REST3 simulation.

2.5 REST3-US simulations

Two sets of REST3 simulations were performed. The first set of REST3-US simulations were conducted to test sampling across different maximum scaling factors (max SF), while the second set of REST3-US simulations were conducted for the calculation of the PMF. In the first set, four REST3 simulations were performed with *maximum* SFs of 0.8, 0.6, 0.4 and 0.2. and the peptide was held at a COM distance of 2.6 nm away from the membrane. The number of replicas was optimised so that exchange rates were approximately 0.20. See Table S1 in the SI for a full description of the replica exchange scaling regimes. Implementing replicas that exhibit an exchange rate of 0.20 balances the need for sufficient energy overlap between replicas for exchanges to occur, whilst keeping the number of replicas to a minimum to reduce

computational resources. Exchanges between replicas were attempted every 50 steps (0.1 ps), unless otherwise stated. REST-US simulations were run for 50 ns.

The second set of REST3-US simulations was used for calculating the PMF. As in the above-described conventional US simulations, windows were spaced at 0.2 nm. For windows with COM-distances > 4.2 nm no interaction scaling was required and the data from the conventional US simulations was used instead. For windows between distances of 4.2 and 2.2 nm (11 windows), a maximum scaling factor of 0.2 was used. Each REST3-US window was simulated for 50 ns. The PMF was reconstructed using WHAM, as described for the conventional US simulations, using the full 50 ns from each window.

2.6 Assessment of configurational sampling

To assess the extent of configurational sampling of the peptide, the Cα atoms of the Cys residues C2, C6 and C11 were used to define three vectors that describe the orientation of the peptide with respect to the membrane normal. These vectors were: 1) C2-Cα - C6-Cα (C2-C6), 2) C6-Cα - C11-Cα (C6-C11), and 3) a vector normal to C2-C6 and C6-C11 (Fig. 3). The Python library MDAnalysis was used to calculate the angle between each vector and the z-axis as a function of simulation time and to obtain frequency distributions for each angle.(76) These distributions were normalised by dividing by the sine of the angle. This method of normalization is herein referred to as 'angle normalised'. After normalisation, the angle distribution for a freely rotating peptide is flat, as each angle is equally likely to be sampled. To assess the effect of scaling, the normalised angle distributions from conventional US and REST3-US simulations for selected windows were compared.

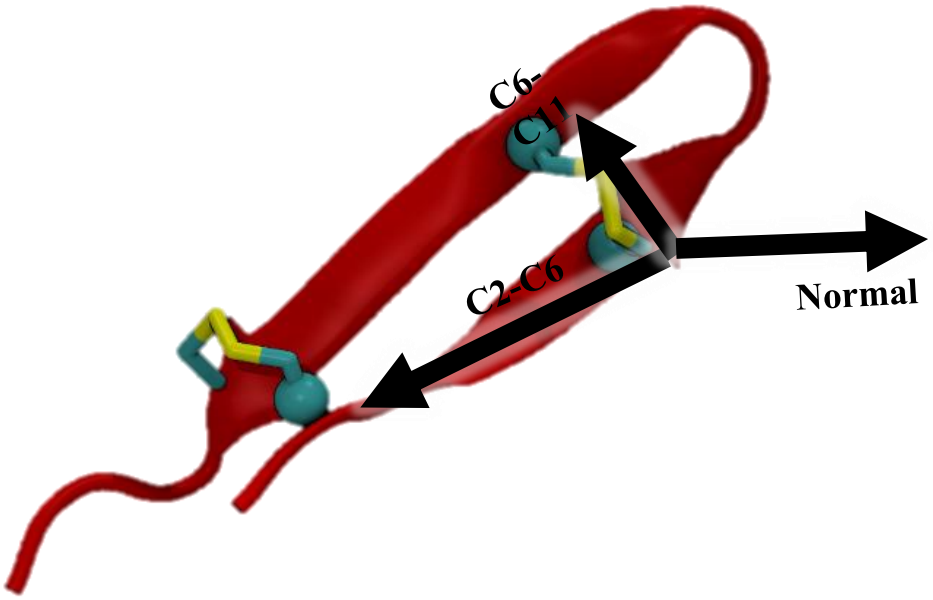


Figure 3: Diagram showing the location and direction of the vectors used to assess rotational motion of the peptide. The vectors use the Ca atoms of the Cys residues. The vectors are from Cys2 to Cys6 (C2-C6), Cys6 to Cys11 (C6-C11), and a vector normal to the plane defined by C2-C6 and C6-C11.

2.7 Hydrogen bond analysis

To calculate the number of hydrogen bonds across time, the Gromacs tool g_hbond was used. An angle cut off of 37° and distance cut off of 0.35 nm was used to define a hydrogen bond. A histogram of the number of hydrogen bonds was plotted using the Python libraries, Seaborn and Matplotlib.(77)

2.8 Calculation of solvation index

To calculate the number of water molecules that are present at the peptide-membrane interface, a Python script utilising the module MDAnalysis was created. The algorithm finds the intersection of all the water molecules within 1 nm of the peptide and within 1 nm of the membrane (i.e. a distance-based criteria), but only for water molecules that are located between the peptide and the membrane (Fig. 4). To find the water molecules located between the peptide and the membrane at any given moment in time during the simulation, three cylindrical projections from the peptide were used. Each cylinder had a width of 0.5 nm and was projected onto the bilayer centre. The height of each of the three cylinders was defined by the COM positions of three groups of residues in Gm: the COM of the entire protein, the COM of residues 1, 16, 17 and 18 (around the C- and N-termini), and residues 7, 8, and 9 (the β-hairpin loop). Three cylinders instead of a single cylinder were used as the cylinder has a fixed radial width.

Testing with one cylinder showed that if a width was chosen that matched the length of the major axis of Gm, then water molecules that are to the side of Gm and not involved in solvation would also be selected. If the width was that of the minor axis of Gm, then water molecules solvating the edges of the peptide would not be counted. Three cylinders are able to roughly approximate the shape of Gm as projected onto the *xy* plane of the membrane. This approach to selecting water molecules was used to intersect with the selection of water molecules around Gm and near the lipid bilayer, to confine the selection to water molecules strictly at the interface. For example, if only the cylindrical projection is used and the peptide had an angle at 45° with respect to the membrane, where only the tip of the peptide may be interacting, a large number of waters would be counted as being in the solvation layer, despite the fact that the presence of these water molecules is due to the orientation of the peptide. This is why the intersection of these two selections is used.

The distance parameters were optimised on the basis that they provided good agreement with the solvation trends visually observed in the unscaled and scaled simulations with SFs of 0.6 and 0.2.

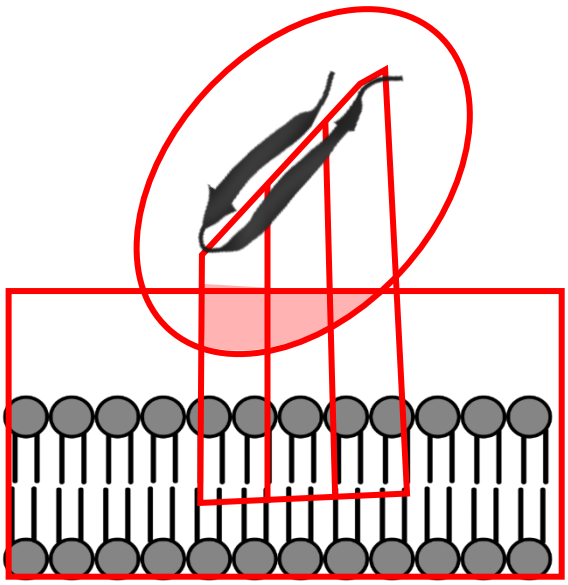


Figure 4: Illustration of the definition of the solvation layer. The red shaded area is designated as the 'solvation layer' area. It is defined by the intersection of the water molecules at a distance of 1 nm around the membrane, the water molecules at a distance of 0.5 nm around the peptide, and the water molecules present in at least one of the three cylindrical selections from the peptide down to the COM of the membrane.

As the cylindrical projection area changes with the orientation of the peptide, the number of water molecules in the solvation layer was divided by an approximation of

the projected selection area. This approximation was calculated by measuring the distance in the *xy* plane between the Cα atoms of residues 9 and 15 (which are at opposite ends of the peptide). This distance was then used as the length of the area, the width of which is assumed to be fixed at 0.5 nm, chosen to approximate the width of the peptide. The number of solvation water molecules divided by the area is termed the 'solvation index'.

3.0 Results and Discussion

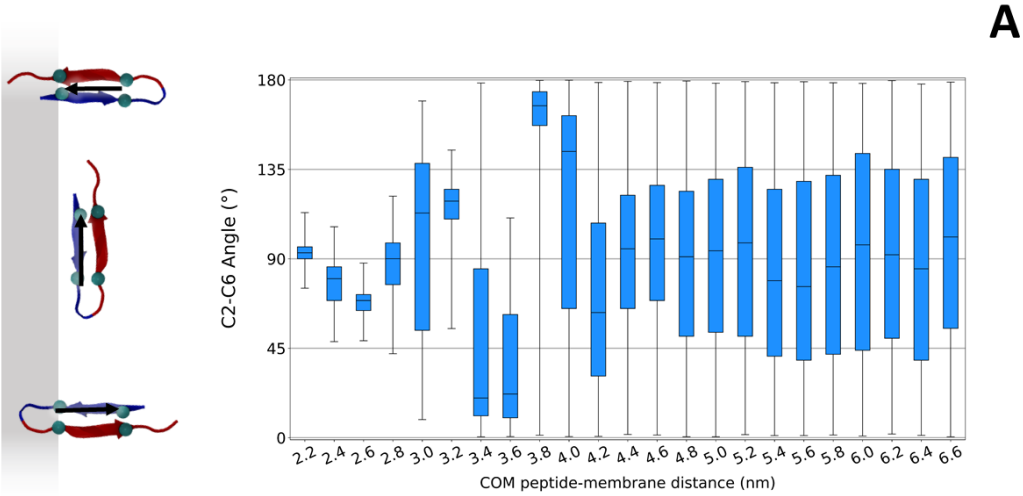
3.1 Determination of windows that require REST3

REST3-US simulations are significantly more computationally expensive than conventional US simulations and it is thus desirable to only use REST3 for windows where the peptide shows reduced rotational motion. To assess at which peptide-membrane distance REST3 is required, the orientation of the peptide with respect to the membrane was defined using the three vectors illustrated in Fig. 3. For each window in the conventional US simulations, the simulation was run for 300 ns, the distribution of the three angles was calculated and normalised, which are shown as boxplots (Fig. 5) (Table 1: simulations cUS-2.2 → cUS-6.6.

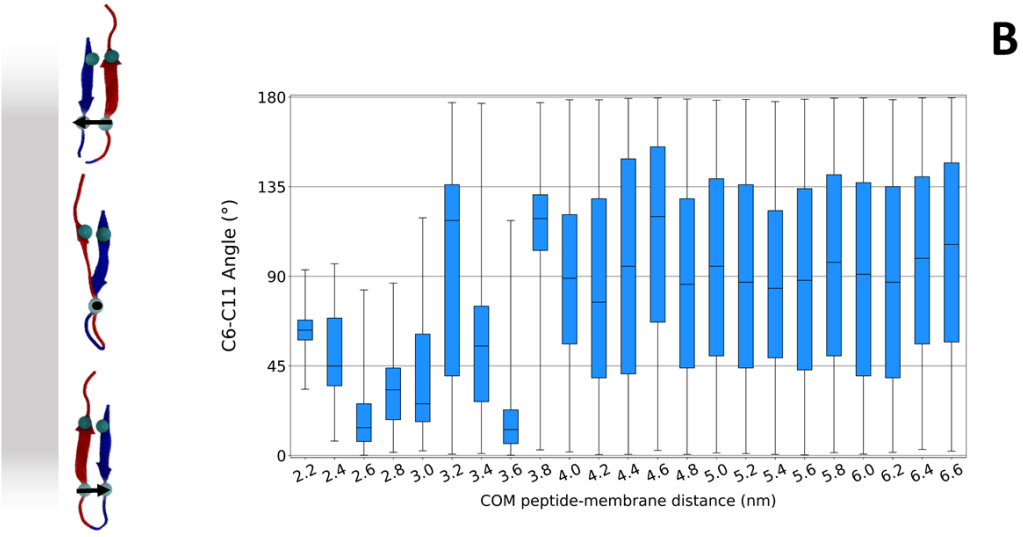
For the C2-C6 vector, the peptide is able to freely rotate at peptide-membrane COM distances greater than 4.6 nm (Fig. 5A). As the peptide continues to approach the membrane, a bias appears towards 180° for distances of 3.8 and 3.9 nm. This occurs as the peptide is strongly attracted to the membrane, and at this distance it can only interact with the membrane if the peptide aligns its major axis (which is approximated by the C2-C6 vector) perpendicular to the membrane. At distances closer than 4.2 nm a narrowing of the angle distributions can be seen. These distributions tend to be around 90°, in which the peptide is parallel to the membrane surface. This is to be expected, as at these distances the peptide would need to be inserted into the membrane to access polar orientations (0° and 180°). For both the C6-C11 and normal vectors, the effect of the membrane on the peptide orientation occurs at similar distances. The angle distributions become narrower after the peptide approaches within a COM-distance of < 4.2 nm. The C6-C11 vector, for distances from 4.0 to 3.2 nm, is not consistently biased towards any angle but, as it gets closer to the membrane, it becomes biased towards 0°, before trending back up to 60° (Fig. 5B). For the normal vector, the angle distributions remain at a value of around 90°, before than trending towards 30° as the peptide is positioned further into the membrane (Fig. 5C). This behaviour

suggests that as Gm turns parallel to membrane, it first resides with one side of its β-hairpin interacting with the membrane (Fig. 5). This stretch of the peptide is where the hydrophobic Val12 and Tyr14 are located, in addition to the small disordered tail of the peptide. As the harmonic potential restrains the peptide even closer to the membrane (< 2.6 nm), the other side of the beta-hairpin turns towards the bilayer, so that the peptide is now flat pressed against the membrane. Based on these observations, it was decided that REST3 would only be used in simulations for windows with a COM-distance < 4.2 nm. In addition to estimating the COM-distance at which REST3 is needed, the narrow distributions of these angles shown in the boxplots in Fig. 5 also give a first indication of the very limited sampling of the rotational motion that is achieved in conventional US simulations.

A


C2-C6 Angle (°)
180
135
90
45
0
2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0 5.2 5.4 5.6 5.8 6.0 6.2 6.4 6.6
COM peptide-membrane distance (nm)

B


C6-C11 Angle (°)
180
135
90
45
0
2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0 5.2 5.4 5.6 5.8 6.0 6.2 6.4 6.6
COM peptide-membrane distance (nm)

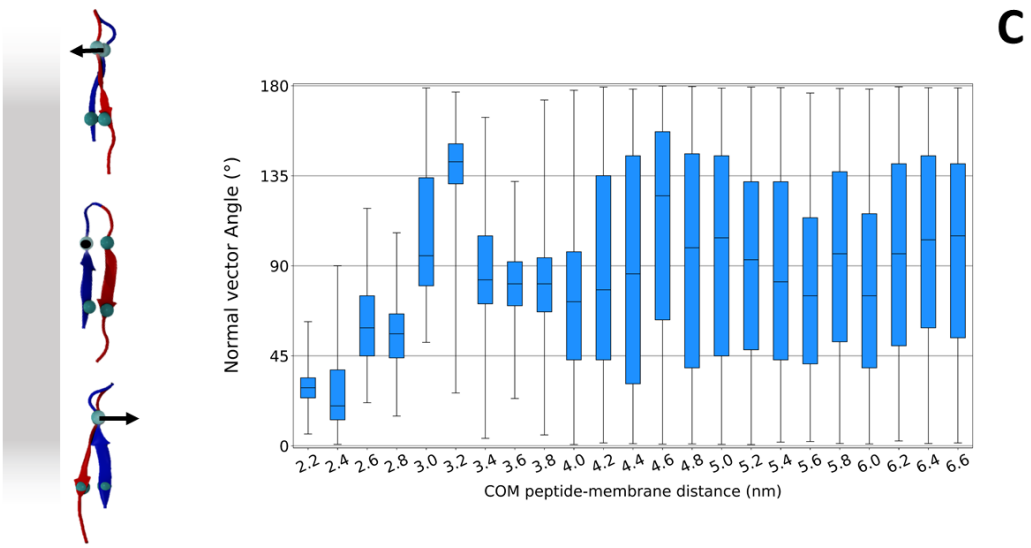


Figure 5: Boxplots of normalised angle distributions across windows in conventional US simulations (see methods section 2.6) (Table 1: simulations cUS-2.2 → cUS-6.6). A: C2-C6. B: C6-C11. C: Normal. A perfectly uniform sampling distribution will have a mean of 90°, an upper quartile range of 135°, a lower quartile range of 45°, and minimum and maximum values of 0° and 180°. Each window was run for 300 ns. Illustrations to the left of the graphs represent the orientations of the peptide at angles of 180°, 90°, and 0° from top to bottom. See Methods section for an explanation of how these angles were calculated.

3.2 Optimisation of scaling factors for REST3

For REST3 to be useful to accurately predict ΔG_b from a PMF, the scaling factors used must increase the rotational sampling of the peptide compared to conventional MD, and thus overcome the restricted motion evidenced in the boxplots in Fig. 5. To determine which SFs will increase rotational motion, a series of simulations with modified Hamiltonians, but with no replica exchange, was performed. These will be referred to as 'scaled simulations'.

In the following analysis of the scaled simulations, the single SF value is the factor by which all peptide-membrane and peptide-water interactions were scaled (collectively referred to as the peptide-environment interactions) (Table 1: simulations cUS-2.6 and SS-0.8 → SS-0.2). These SFs ranged from 1.0 (unscaled) through to 0.8, 0.6, 0.4 and 0.2 (a total of five simulations). For example, a scaled simulation with a SF of 0.6 means that all peptide-environment interactions (both vdW and electrostatic interactions) are scaled down using a factor of 0.6 (i.e. their magnitude was 60% of their full value). All scaled simulations were performed at a COM-distance of 2.6 nm, where the peptide is at the membrane surface. At this distance, the peptide has

restricted rotational motion (Fig. 5) for all three orientation vectors. Each simulation was run for 200 ns and the three vectors used previously for the boxplots were used to assess sampling of the rotational motion. However, in this analysis the distributions of the three angles were normalised, as described in the Methods, and are shown as histograms (normalised frequency vs angle). The angle histograms for each of the three vectors from each of the five scaled simulations are shown in Figs. 6A, B and C. As the angle histograms are angle normalised, complete rotational freedom of the peptide is indicated by a flat distribution (i.e. all angles are equally likely). Any increase in frequency in certain angles indicates the occurrence of bias towards a specific orientation of the peptide on the membrane. The aim of the scaling simulations was to find a SF that enables the peptide to access as great a range of orientations as possible. At the same time, the further away from a value of 1.0 that the SF is, the more replicas will later be required to obtain an appropriate exchange rate during the full REST3 simulations. Hence, a SF that is as close as possible to 1.0 whilst still enhancing the sampling is sought.

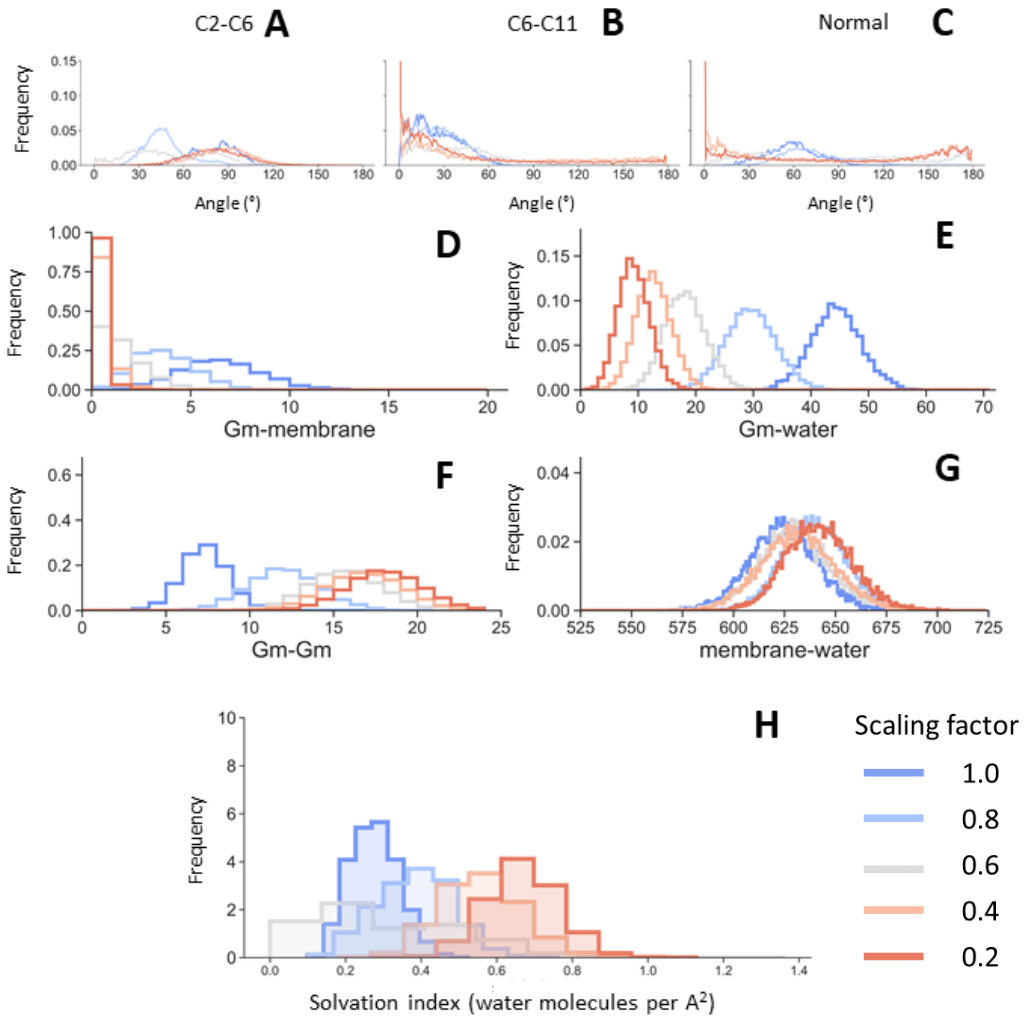


Figure 6: Summary of sampling and interactions in scaled simulations of peptide-environment SF of 1.0, 0.8, 0.6, 0.4, and 0.2 (Table 1: simulations cUS-2.6 and SS-0.8 → SS-0.2). In these simulations the peptide is held at a COM distance of 2.6 nm to the membrane. A-C: Frequency and angle normalised distributions across different scaling factors (see Methods section 2.6). Hotter (red) colours represent a greater degree of scaling. From left to right, the distributions of the C2-C6 vector, C6-C11 vector, and the normal vector. D-G: Histograms of the number of hydrogen bonds between system components, for D: Gm-membrane, E: Gm-water, F: Gm-Gm, and G: membrane-water. H: Histograms of solvation index. The high values at an angle of 0° are an artifact of the angle normalisation.

From the angle histograms shown in Fig. 6B it can be seen that for vector C6-C11, in the unscaled simulation (SF of 1.0) the motion of Gm is restricted to a range of 0 to 75°, with most sampling around 20-30°. Scaling interactions down to 0.8 results in minimal changes in the angle distribution. A SF of 0.6 has a broader distribution, ranging from 0 to 100°. At a SF of 0.4, Gm is able to sample all angles, but still exhibits

a bias towards 0°. A SF of 0.2 results in a similar range but orientations are slightly more biased towards 0° compared to simulations with a SF of 0.4. Sampling of orientations of the normal vector is similar to that of C6-C11 (Fig. 6C). In the unscaled simulation (SF of 1.0) the vector has a range of angles between 25 and 100°. There is minimal change upon scaling by 0.8. At a SF of 0.6, however, the vector is able to access all possible angles, as indicated by the almost flat distribution. The same is true for SFs of 0.4 and 0.2. For the C2-C6 vector (Fig. 6A), the change of angle distribution across scaling factors is markedly different in comparison to the other two vectors. In the unscaled simulation, the vector mostly samples angles from 50 to 100°. The distribution of angles for a SF of 0.8 is shaped differently to that of 1.0, but still suggests a comparable range of sampling from 25 to 90°. The broadest of distributions occurs with a SF of 0.6, ranging from 0 to 100°. Both SFs of 0.4 and 0.2 result in sampling of angles between 50 and 125°. In summary, these results suggest that for the C6-C11 and normal vectors, a scaling factor of 0.4 is sufficient to for all orientation angles to be sampled even if there is a small bias left in the case of the C6-C11 vector. However, for the C2-C6 vector, no SF tested appears to allow the peptide full rotational movement along that axis. Unexpectedly, decreasing the factors for this vector does not seem to directly relate to how broad the sampled angle distribution is: further scaling down of the interactions does not necessarily improve the sampling. Use of a SF of 0.6 leads to sampling of a wider range of angles than with SFs of 0.4 and 0.2, and there is little change in orientations sampled between simulations using SFs of 0.4 and 0.2, where further improvement of sampling would be expected.

The reason for this difference in the sampling of angle distributions across SFs for the different vectors might be related to the shape of Gm. As seen in Fig. 2, Gomesin can be roughly represented as being ovoid in shape. The C2-C6 vector runs along its major axis while C6-C11 and the normal represent the motion around the minor axis. Around the minor axis, the shape of the peptide is relatively radially symmetrical, and so motion around this axis will have a relatively consistent steric hindrance against the membrane. But around the major axis or the C2-C6 vector, at polar angles (0° and 180°) the peptide will be extended into the membrane at short distances, while at 90° the peptide will lie flat on the membrane. Even though the peptide-membrane interactions have been reduced, steric hindrances would provide a large energetic barrier and prevent the peptide from sampling polar angles along that axis. Thus, the C2-C6 vector will have greater difficulty sampling polar angles than the other two

vectors and, consequently, there should not be the expectation that the angle distributions of the C2-C6 vector will continue to broaden continuously as the interactions of the peptide are reduced with lower SFs. However, it remains unclear why at an SF of 0.6 the C2-C6 vector of Gm is able to sample a wider distribution of angles compared with the lower SFs of 0.4 and 0.2. An understanding of how the value of the SF changes the balance of interactions between the different components of the system is thus needed.

3.3 Effect of scaling on the interactions in the system

There is likely to exist competition for "binding sites" in a peptide-membrane system such as the one studied here. For example, both the peptide and the water molecules are attracted to the lipid head groups in the membrane. For the peptide to bind to the membrane (i.e. transition from being in solution to being membrane-bound), it must form peptide-membrane interactions that replace the peptide-water and membrane-water interactions. Consequently, an appropriate area of the membrane and a corresponding area of the peptide need to be desolvated. When the interaction between two components of the system is scaled down, interactions with another unscaled component may become more dominant.

To shed light on this, Figs. 6D, E, F, and G show the frequency histograms of the number of hydrogen bonds (HBs) between the peptide, water and membrane components of the system for the same SFs whose angle distributions were presented in the previous section. As would be expected when the peptide-membrane and peptide-water interactions are scaled down, the number of Gm-membrane and Gm-water HBs is progressively and substantially reduced with decreasing SFs (Fig. 6D, E). Interestingly, in the unscaled peptide-peptide and membrane-water interactions, the number of Gm-Gm and membrane-water HBs also exhibit significant changes (There are, however, no observable changes in the number of water-water HBs as the SF is reduced see Fig. S1 for the corresponding histograms). The intra-peptide HBs increase as the scaling factor decreases (Fig. 6F). In the unscaled simulation, Gm has a median number of internal HBs of 7 ± 1.4. As the SF is reduced the median number of HBs increases to 12 ± 2.2, 15 ± 2.2, 16 ± 2.3 and, finally, to 18 ± 2.2 when the SF is 0.2. This can be rationalised by understanding that the magnitude of the interactions of the peptide with its environment are reduced, leading to a likely imbalance in the force field terms that manifests itself as an increase in intra-peptide HBs. This results in the peptide 'curling' inwards and its secondary structure being lost (Fig. S2 in the

Supporting Information) Likewise, the number of membrane-water HBs increases as the SF decreases (Fig. 6G). The unscaled simulation has a median of 624 ± 16.5 membrane-water HBs. When the system is scaled by 0.8, the median number of HBs increases by around a standard deviation to a value of 638 ± 15.0. However, this value reduces when the system is scaled by 0.6 and 0.4, with median values of 629 ± 17.8 and 631 ± 17.9 HBs, respectively. This value then increases as the system is scaled by 0.2 to 641 ± 16.3 HBs. As the peptide has reduced interactions with water and the membrane, it could be expected that it can no longer effectively displace water molecules from the surface of the membrane, resulting in water preferentially interacting with the membrane. Hence, for values of the SF from 1.0 to 0.2, there is an increase in the number of membrane-water HBs but it does not follow a direct trend and the differences between simulations is slight, as the water molecules involved in the solvation of the peptide-membrane interface are only a small percentage of all the water molecules that interact with the membrane.

Computation of the solvation index was used to further characterise the behaviour of water at the interface of the membrane and peptide. Fig. 6H shows the histograms of frequency vs solvation index across for simulations run with different SF. The solvation index is directly related to the number of water molecules per unit area at the peptide-membrane interface and hence the density of the solvation layer. In the unscaled simulation, a median of 0.27 ± 0.06 water molecules per A^2 of interfacial area was computed. A SF of 0.8 increases the solvation of the peptide-membrane interface to 0.38 ± 0.10 water molecules per A^2. A SF of 0.6 does not fit with the trend of the other scaling factors, having a median of 0.25 ± 0.19 water molecules per A^2 and a much broader range of solvation indices accessed. This unusual behaviour is not continued for both SFs 0.4 and 0.2, with median solvation indices of 0.55 ± 0.11 and 0.66 ± 0.10, respectively. At these SFs, the peptide exhibits the most configurational freedom; however, this is accompanied by a more than doubling of the number of water molecules at the peptide-membrane interface. This analysis better highlights the change in relative membrane binding affinities, and the increase in solvation corresponds to the order of the scaling factor, with the notable exception of SF = 0.6. At a scaling factor of 0.6, there is a broad range of solvation indices reached in the simulation, indicating large changes in the peptide-membrane solvation has occurred (Fig. 6H). The peptide-membrane interface is even frequently less solvated than what is observed in the unscaled simulation. Visualization of the simulation with SF = 0.6

reveals that the lipid head groups move out from under the peptide (Fig. 8). This behaviour can be explained by the fact that upon scaling, the membrane has more favourable interactions with water than the peptide does. To maximise interactions, water molecules thus displace the peptide from the membrane or, if the peptide cannot be displaced (as indeed occurs in these simulations), lipid molecules move out from underneath the peptide to better interact with water. This effect is more apparent when testing scaling factors at COM distances that are closer than 2.6 nm, as water molecules need to displace the peptide further, which becomes increasingly difficult due to the harmonic restraint imposed. For example, at 2.2 nm, for all scaling factors tested, the peptide inserted into the membrane – orienting towards the polar angles of the C2-C6 vector, so as to minimise the force required to displace lipid molecules (Fig. S3 in the Supporting Information). Therefore, the broadened C2-C6 angle distribution in the SF = 0.6 simulation can be explained by the fact that lipid molecules move to minimise the area that does not interact with the peptide and to maximise their interactions with water. This in turns allows the peptide to access angles that would have caused steric clashes if the lipid molecules had not shifted. Consequently, as the peptide shifts from a parallel to a polar orientation, there is a broadening of the C2-C6 angle distribution. As the lipids move out from under the peptide to search for interactions with water molecules, the solvation index of the peptide is reduced for SF = 0.6. This effect is likely random and has occurred at a SF of 0.6 due to chance.

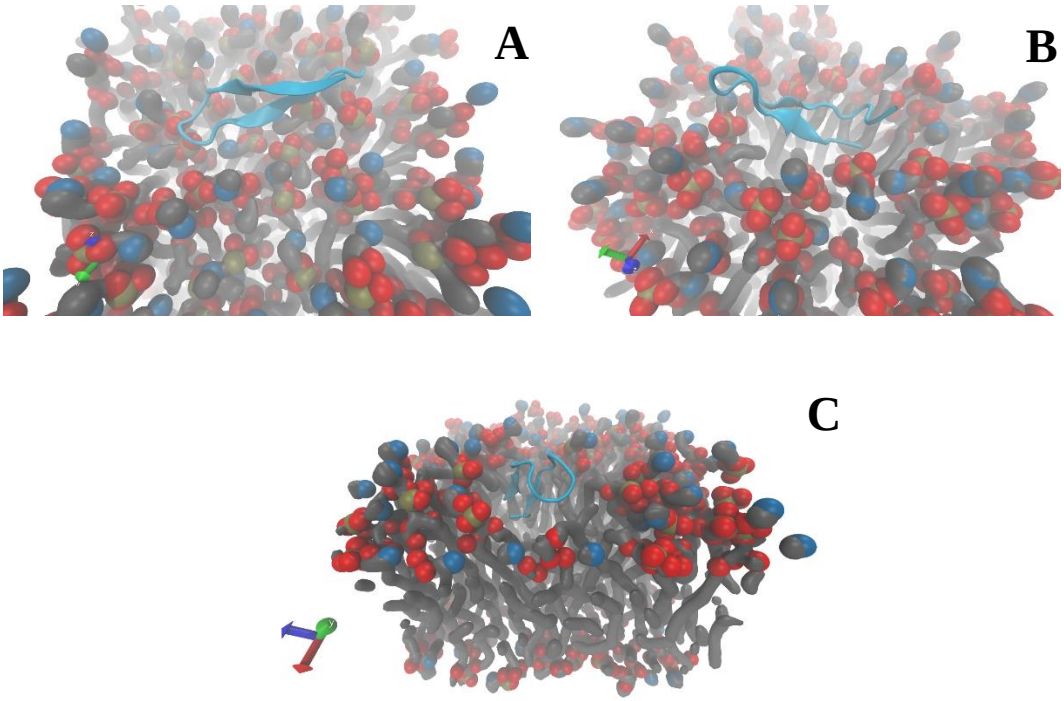


Figure 7: Insertion of Gm into the bilayer caused by the scaling of peptide-environment interactions. Lipids are coloured according to atom type: grey-carbon, red-oxygen, blue-nitrogen and gold-phosphate. A: Snapshot of an unscaled simulation, where Gm is held at a COM distance of 2.8 nm (Table 1: cUS-2.8). This image is included as a reference for normal peptide-membrane interface behaviour. B: Snapshot of the scaled simulation with a peptide-environment interaction SF of 0.6 (Table 1: SS-0.6). Here the lipid headgroups have moved out from under the peptide and the lipid tails are exposed to the peptide. C: Snapshot of a scaled simulation where the peptide-environment interaction SF is 0.2 and the peptide is held at a COM distance of 2.2 nm to the membrane (Table 1: SS2.2-0.2). Here the effect is more severe.

In summary, scaled simulations and subsequent analysis of the interactions in the system showed that sufficient configurational sampling can be achieved with a SF of 0.4. Selective scaling of the interactions in the system was observed to cause changes in the interactions between other components (molecules) in the system that had not been scaled. The relative strength of membrane-water and peptide-membrane interactions changes such that the membrane will attempt to minimise its interactions with the peptide. This results in either water displacing the peptide from the membrane or the peptide inserting into the membrane so that the head groups can regain interactions with water. As the peptide reduces its intermolecular interactions upon scaling, it gains internal hydrogen bonds, which results in changes to the structure.

3.4 Effect of maximum scaling factor on replica exchange simulations

To test the SF used in the scaling only simulations (Table 1: SS-0.8 → SS-0.2) when incorporated into in a replica exchange regime, REST3 simulations with maximum

SFs (Max SF) of 0.8, 0.6, 0.4 and 0.2 were performed at a COM-distance of 2.6 nm, at which there is a minimum in the conventional US PMF, as discussed further below (Fig. 9) (Table 1: simulations cUS-2.6 and R-0.8 → R-0.2). For the different Max SFs tested, a different number of replicas were used with each regime, ranging from 3 for Max SF 0.8 to 10 replica for Max SF 0.2. The SFs of replicas in any particular regime were generally spaced at intervals of 0.1 (Table S1 in the Supporting Information). This resulted in exchange rates of around 20%. As each of the regimes has a different number of replicas, the total time simulated across all replicas is different in each regime. To assess the impact of increased simulation time on the sampling and compare each regime to the equivalent amount of simulation time from unscaled simulations, 10 separate unscaled simulations were run for 50 ns each, all starting from the same starting structure. The sampling data was combined from 1, 3, 5, 7 and 10 simulations (to match the number of replicas in each of the REST3 regimes). The additional data does not significantly change the configurational space sampled by Gm (Fig. S4A-C in the Supporting Information). Therefore, the effect of increased simulation time between replica exchange regimes has been controlled for and, hence, the observed differences between regimes is the result of the use of different Max SF. Fig. 8 shows the histograms that describe the orientational sampling of the peptide, its interactions between the other system components and the interfacial solvation of the peptide. In Fig. 8, the unscaled data reported is from the combination of 10 unscaled replicate simulations.

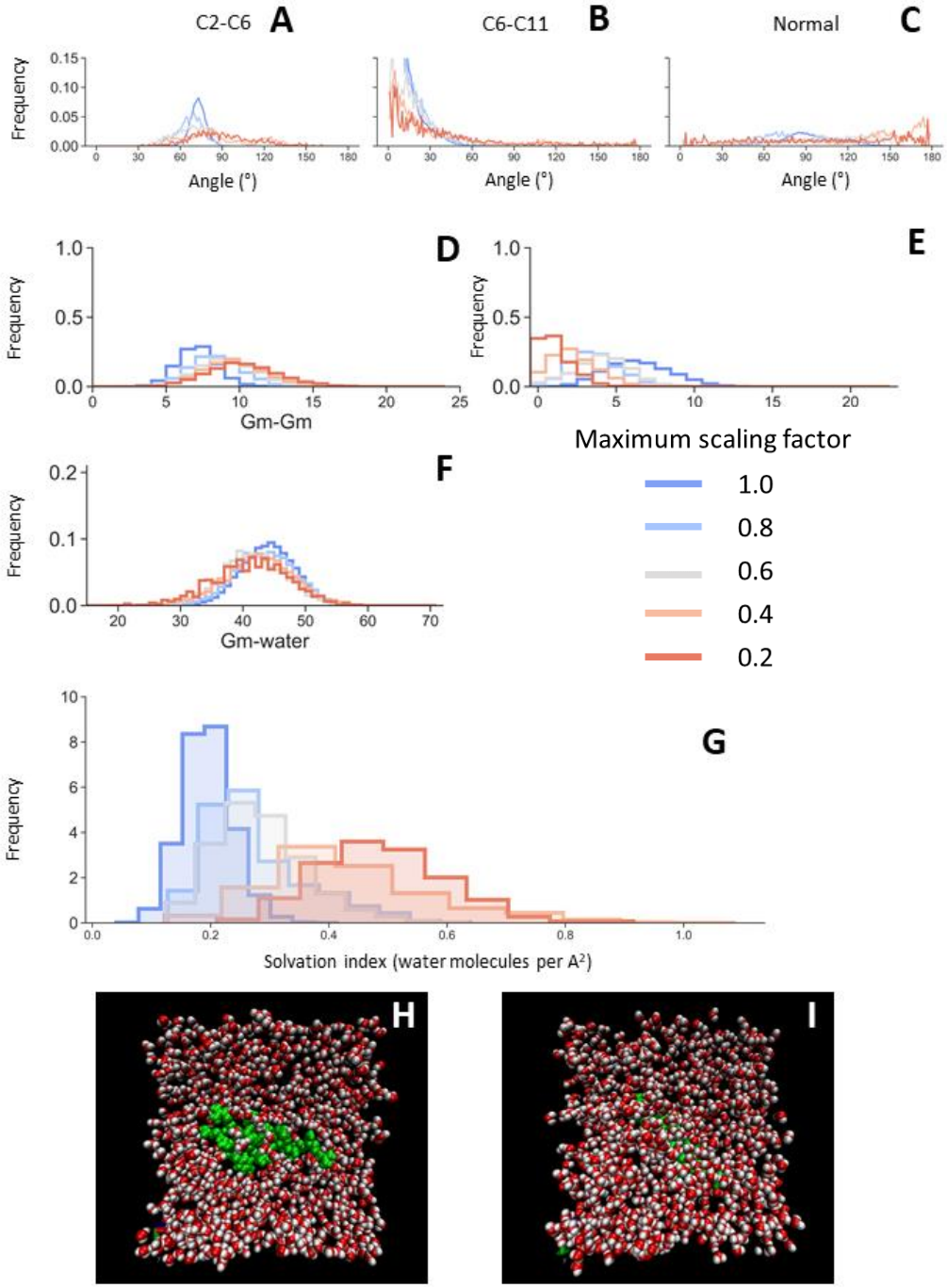


Figure 8: Summary of tests of sampling and interactions in maximum scaling factor REST3 simulations with maximum peptide-environment SFs of 0.8, 0.6, 0.4, and 0.2 (Table 1: R-0.8 → R-0.2). In these simulations the peptide is held at a COM distance of 2.6 nm to the membrane A-C: Frequency and angle normalised distributions across different scaling factors (see Methods section 2.6). From left to right: distributions of the C2-C6 vector, C6-C11 vector and the normal vector. Graphs have been truncated at a frequency of 0.15 for visualisation purposes. D-F: Histograms of the number of hydrogen bonds between system components: Gm-Gm, Gm-membrane, and Gm-water. G: Histograms of solvation index. H: Snapshot from

the conventional MD simulation of the water layer observed from the membrane. Gm is shown in green and can be seen breaking through the water molecules at the interface. I: Snapshot from the ground replica in REST3 (with a max SF of 0.2) of the water layer observed from the membrane.

In Fig. 8A, for the C2-C6 vector the angle distribution for the unscaled simulation (SF = 1.0) is strongly peaked, with a range of only 50° to 90°. Max SFs of 0.8 and 0.6 exhibited the same upper range as the unscaled simulation, but Gm can access slightly lower angles, with the lower range of the angle distribution almost reaching 25°. For Max SFs of 0.4 and 0.2, like in the scaled simulations, the broadest range of angles are sampled, going from around 50° to 125°. The angle distributions of the C6-C11 vector are similar between the scaled and the replica exchange simulations (Fig. 8B and Fig. 6B, respectively). In the unscaled simulation, the peptide exhibits a strong preference for orientations towards 0°. This does not improve with a Max SF of 0.8. A Max SF of 0.6 reduces this bias for 0°, but a significant change only occurs when the replica exchange regime has a Max SF of at least 0.4. Both the C6-C11 angle distributions for Max SFs 0.4 and 0.2 still exhibit a large bias near 0°, but their ranges extend all the way to 180°. In the unscaled simulation, the normal vector can access a wide range of angles from 25° to 140°. A Max SF of 0.8 does not have a significant impact on this angle distribution, but an improvement is seen with Max SF of 0.6. Its upper range is around 150° and extends down to 0°. Simulations with Max SF 0.4 do not reach 0° but rather reach a lower range just below 25°. This range, however, extends all the way to 180°, where there is a slight peak. A Max SF of 0.2 results in the widest sampling, as the normal vector can sample from 0° to 180°.

Overall, the trend in sampling across scaling factors observed is similar between the scaled and REST3 simulations (simulation sets SS and R), indicating that the orientations of Gm found in the scaled replicas are being exchanged down into the ground replica, and increasing the sampling. The exception to this is Max SF = 0.6, where the unusually broad distribution of the C2-C6 angle that was seen when scaling the system at 0.6 is no longer observed, and the peptide does not appear to become trapped in the membrane (Fig. 6A). In terms of overall sampling of orientations, use of both Max SFs of 0.4 and 0.2 lead to similar angle distributions. Use of Max SF = 0.2 leads, however, to C6-C11 angle distributions marginally less biased towards 0° (Fig. 8B) and with a slightly broader normal angle distribution (Fig. 8C). Use of a Max

SF of 0.2 requires the use of 10 replicas, as compared to 7 replicas in Max SF = 0.4. Due to the slightly increased sampling of the C6-C11 and normal angles, a Max SF of 0.2 was chosen to be used for the subsequent calculation of the entire PMF using REST3. It was not possible to determine at this point whether the slight improvement in sampling was worth the increased computational cost.

The HBs between the system components in the REST3 (set R) simulations are observed to change proportionally to the Max SFs used (Fig. 6D-F,). Figs. 6D-F show histograms of the number of HBs between different components. Data for the unscaled simulation is combined over 10 unscaled simulations for reasons described previously (Fig. S4D-H in the Supporting Information). The disparity in the number of HBs between the unscaled simulations and the REST3 ground replica is smaller than between the unscaled and scaled simulations (Fig. 6). This demonstrates that interactions are regained as the configurations move towards the ground replica. There are, however, differences in the HB pattern that are in proportion to the Max SF.

In comparison to the unscaled simulations, all the REST3 simulations still exhibit an elevated number of Gm-Gm HBs (Fig. 8D). In the unscaled simulations, Gm has a median of 7.0 ± 1.4 intramolecular HBs. As seen in the scaled simulations, in the REST3 simulations the median number of intramolecular HBs increases with decreasing Max SF, going to 8.0 ± 1.7, 9.0 ± 2.0, 9.0 ± 2.2, and finally 10.0 ± 2.4 for Max SF = 0.2. The difference of 3 HBs between the unscaled and REST simulation with a Max SF of 0.2 is much smaller than the difference of 11 HBs between the unscaled and the scaled simulation at an SF of 0.2. Though there is some decrease in the median number of HBs predicted in the ground replicas of the REST3 simulations, the differences suggest that Gm still experiences structural changes as in the higher replicas. Visual inspection of the structure also corroborated this, showing similar structural disturbances to those seen in the scaled simulations (Fig. S5 in the Supporting Information). Circular dichroism (CD) experiments of Gm in the presence of the single-tail lipid lysophosphatidylcholine showed that its structure is largely unaffected in a membrane-like environment.(21) It should then be expected that in all of the simulations Gm should retain its rigid β-hairpin structure, which has been characterised by NMR spectroscopy.(37, 62) The conformational changes observed in Gm in the simulations are thus not associated with membrane binding, as is indeed the case in some AMPs, but an artefact of the REST3 simulation due to the SFs used.

The number of Gm-membrane HBs is reduced in the REST3 simulations compared to the unscaled simulation (Fig. 8E). Across the 10 unscaled replicate simulations, the median number of Gm-membrane HBs was computed to be 6.0 ± 2.1. In the REST3 simulations this value decreased to 4.0 ± 1.6, 4.0 ± 1.9, 2.0 ± 1.5 and 1.0 ± 1.1 as the value of Max SF decreased. Due to the use of REST3, there is the expectation that higher energy states will be sampled in the higher scaled replicas, some of which will end up being present in the ground replica. These higher energy configurations would likely have reduced Gm-membrane HBs (as these are low-energy, favourable interactions). However, it should still be expected that these low energy states be sampled in the ground replica. Given that states with a large number of Gm-membrane HBs (> 6) are not sampled in the Max SF 0.4 and 0.2 regimes compared to the unscaled simulations, and that the Gm-membrane interactions are explicitly reduced in the non-ground replicas, it is suspected that in the ground replica the system does not regain these interactions at a fast-enough rate for high numbers of HBs to be observed. Hence, the change in Gm-membrane HBs observed are not only the result of enhanced sampling, but also of insufficient time to allow the interactions to return to 'normal' after the scaling of peptide-membrane and peptide-water interactions.

Fig. 8F also suggests that for Gm-water HBs, there does appear to be slight shift towards fewer peptide-water HBs for reduced Max SFs. The differences are not likely to be significant, at least not in relation to other interactions which show Max SF-related changes. Membrane-water and water-water HB distributions are reported in Fig. S6 in the Supporting Information. Plots of membrane-water HBs were not included in the main text as the solvation index analysis is considered superior at analysing the properties of the solvation layer.

In the scaled simulations, a layer of water was observed to form at the peptide-membrane interface as result of the membrane compensating for lost interactions with the peptide with interactions with water. To assess whether the solvation layer was still present in the ground replica of the REST3 regimes, the solvation index was calculated (Fig. 8G). The water layer was observed to persist in the ground replica (Fig. 8H-I). In the unscaled simulation, there is a median of 0.19 ± 0.04 water molecules per A^2. This is observed to increase as the Max SF decreases, with Max SF 0.8 having a median solvation index of 0.25 ± 0.09, and for Max SF 0.6 and 0.4 the values increasing respectively to 0.28 ± 0.08 and 0.41 ± 0.14 water molecules per A^2. The most solvated condition was observed with a Max SF of 0.2 with a median solvation index of 0.48 ±

0.11 water molecules per A^2. The regimes with the most sampling, max SF 0.4 and 0.2, also had more than double the number of water molecules at the peptide-membrane interface compared to the unscaled system. These extraneous water molecules appear to block interactions between the peptide and membrane and constitute the reason for the reduced number of Gm-membrane HBs. The abnormal solvation trend for a SF of 0.6 is not present in the ground replica with a Max SF of 0.6. This may suggest that the embedding of the peptide observed in the scaled simulations was an unusual event and did not arise by the use of a SF of 0.6 specifically.

In summary, the REST3 method was able to enhance the sampling in the ground replica (Figs. 8A, B and C). For the Gm-POPC system, this was best achieved by using a Max SF of 0.4 or 0.2. Both Max SFs, however, caused changes in interactions between system components. Gm-Gm HBs were observed to increase, which changed the peptide's structure from its rigid beta-hairpin. The number of Gm-membrane HBs was reduced in the ground replicas of both Max SF 0.4 and 0.2 regimes, suggesting that the peptide could not regain the interactions it lost upon scaling of the peptide-environment interactions. The water layer between peptide and membrane was also found to still remain in the ground replicas.

3.5 Potential mean force calculation

PMFs for the binding of Gm to POPC membranes were calculated using conventional US and REST3-US (Fig. 9) (Table 1: simulations cUS-2.2 → cUS-6.6 and RUS-2.2 → RUS-4.2). In the conventional (unscaled) US, the free energy begins to decrease at a COM-distance of 4.2 nm as the peptide approaches the membrane, which corresponds approximately to the distance where the rotation of the peptide becomes biased towards particular orientations (Fig. 5). The free energy minimum of the unscaled PMF is between 2.4 and 2.6 nm, which corresponds to the distance at which the peptide becomes embedded in the lipid head groups and partially into the hydrophobic core of the membrane. The ΔG_b was calculated as the difference in the value of the free energy far away from the bilayer and the absolute minimum. The ΔG_b calculated from the conventional US is -60.0 ± 5 kJ mol^{-1}, which corresponds to a K_D of 0.3 pM. This would suggest that Gm is a very strong binder to POPC. Whilst SPR studies have been performed with Gm to POPC membranes, K_D values were either not reported (44), or, the model fit was poor and accurate K_D values could not be obtained (78). There is, however, an experimental K_D value of 8 µM for the binding

of Gm to a POPC / POPG (70:30) bilayer.(53) POPG is a negatively charged lipid and is used in models of bacterial membranes.(79) The difference in charge between bacterial (POPG-containing) membranes and overall more neutral mammalian membranes is widely thought to be responsible for the ability of AMPs, including Gm, to preferentially target bacterial cells over healthy mammalian cells.(3) Given that the value of our calculated K_D suggests that Gm binds to POPC (a neutral lipid used to model mammalian cells) with an affinity several orders of magnitude stronger than what Gm binds to a POPC/POPG (70:30) membrane, it is likely that the ΔG_b has been severely overestimated by the conventional US simulation.

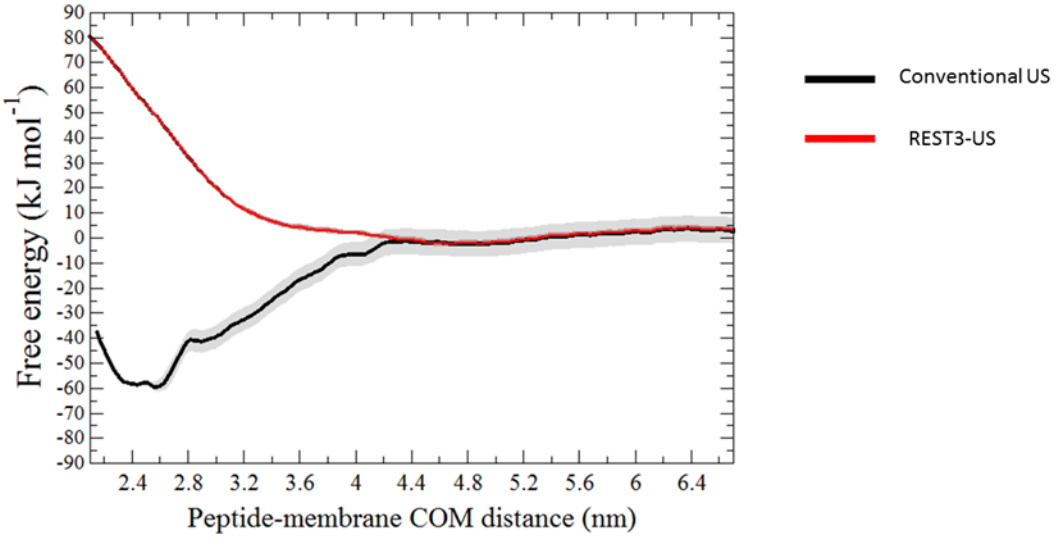


Figure 9: PMF of the conventional US and REST3-US simulations (Table 1: simulations cUS-2.2 → cUS-6.6 and RUS-2.2 → RUS-4.2). The shaded area represents the error, which was calculated using bootstrapping. Note that the error for the REST3-US are too small to see. The REST-US simulations used a max peptide-environment SF of 0.2 and was used for windows 2.2 to 4.2 nm.

Overestimation of ΔG_b in the conventional US simulation is likely to be the result of insufficient sampling. As seen in Fig. 4, Gm samples only a subset of possible configurations, suggesting that the sampled ensemble is likely mostly comprised of lower energy states from a single local minimum. As higher energy states are not visited as often within the time span of conventional MD simulations and the crossing of large energy barriers are infrequent events, there is an overestimation of the ΔG_b. This sampling problem is a well documented problem in MD simulation and is the very reason enhanced sampling methods such as REST3 are required to obtain accurate PMFs for peptide-membrane systems.(80, 81)

The force field parameters for the PC lipid headgroups may also contribute to the overestimation of ΔG_b. Recently published work by Marzuoli et al., (2019) suggests that the partial charges in the choline and phosphate groups may be overestimated in the GROMOS54a7 force field (used in this work).(82) In the study by Marzuoli et al, the headgroups were reparameterised so that the partial charges were more evenly distributed across the headgroup atoms. The new parameters were then used in simulations of binding of the anti-microbial peptide lactoferricin to a POPC membrane. Lactoferricin, like Gomesin, is an amphipathic β-hairpin peptide.(83) From a visual assessment of their trajectories, the authors found that lactoferricin bound more weakly to POPC and was able to more freely re-orient when using the updated parameters. Weaker binding of cationic peptides to POPC membranes is indeed consistent with the experimentally expected behaviour of lactoferricin.(82). However, the authors did not validate these new parameters with free energy calculations to compare predicted binding affinities to experimental data. The updated lipid headgroup parameters would likely improve the accuracy of the estimated ΔG_b of the binding of Gm to a POPC membrane, in addition to allowing increased configurational sampling. In future studies of Gm, simulations using the revised force field parameters should be conducted and compared to the ones reported here.

Comparison of the PMFs produced using the conventional US and REST3-US methods shows that there are striking differences in the behaviour of Gm upon binding to a POPC bilayer. In the PMF predicted by the REST3-US simulation there is no energy minimum and the free energy unexpectedly rises sharply as the peptide approaches the membrane, suggesting that there is strong repulsion between Gm and the membrane. If Gm indeed binds weakly to POPC, as discussed above, there should still be a shallow free energy minimum where the positively charged peptide interacts with the zwitterionic lipid head groups. Whilst the conventional US simulation overestimates the ΔG_b, the disparity between the shapes of the PMFs suggests that the REST3-US simulation significantly underestimates the favourable peptide-membrane interactions.

The repulsion predicted by the REST3-US simulation can be rationalised by investigating how the interactions differ between the methods using the data in Fig. 8. At a COM-distance of 2.6 nm, the peptide is at a free energy minimum in the conventional US simulation, whilst in the REST3-US one there exists repulsion between the peptide and the membrane. Visual inspection and careful analyses of the

simulation revealed that favourable interactions between the peptide and membrane were likely being blocked by the presence of extraneous water molecules residing at the peptide-membrane interface (Fig. 8G). This layer of water may appear to push the peptide away from the membrane surface and create resistance to the movement of the peptide towards the membrane. As noted above, the water layer forms because the scaling of the peptide environment changes the balance of interactions such that the membrane preferentially interacts with water molecules over the peptide. As the water layer is still observed in the unscaled ground replica, this indicates that there is not sufficient time for the peptide to regain interactions with the membrane during replica exchange. Steric clashes at small peptide-membrane COM-distances appear to not be compensated for by the formation of favourable HBs or electrostatic interactions when the water layer is present. Thus, there is a predominance of unfavourable interactions that result in an increase in the free energy as the peptide approaches the membrane.

3.6 Determination of the desolvation timescale of Gomesin

Since the peptide does not displace the water layer and regain HBs with the membrane in the ground replica of the Max SF 0.2 regime, the feasibility of increasing the exchange attempt frequency (EAF) to resolve this problem was investigated. It was hypothesised that increasing the EAF would create greater periods of time between exchanges, which may allow the peptide to regain interactions with the membrane and the extraneous water molecules to be displaced from the peptide-membrane interface. To determine an appropriate EAF, the time needed for the solvation layer to be dispersed by the peptide in an unscaled simulation with no replica exchanges was estimated. Five of these unscaled simulations were performed, where each was started with a different random structure, obtained at random time points from the ground replica of the Max SF 0.2 simulation (Table 1: CC-1 → CC-5). The five simulations were all run for at least 75 ns each. The evolution of the solvation index across time for each of the five simulations was plotted to assess the time needed for the peptide to eliminate the water layer (Fig. 10A).

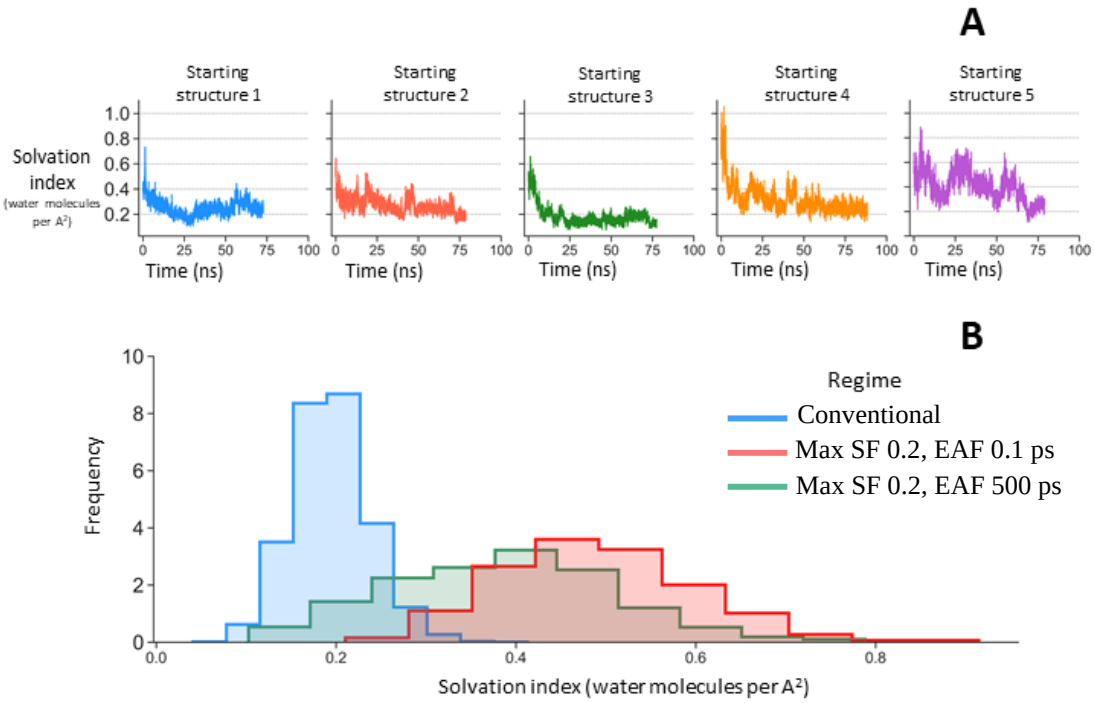


Figure 10: A: Evolution of the solvation index across time for five different starting structures in conventional MD (Table 1: CC-1 → CC-5). The structures were randomly chosen from a REST3 max SF 0.2 ground replica exchange simulation. In these simulations the peptide is held at a COM distance of 2.6 nm to the membrane B: Histograms of the solvation index for the combined data of 10 conventional replicates, REST3 max SF 0.2 ground replicas with an EAF of 0.1 ps and 500 ps (Table 1: simulation set Cr, R-0.2, and R-EAF). Replica exchange simulations were run for 50 ns.

Each of the simulations shows a downwards trend in the solvation index, indicating that a reduction in the number of water molecules at the peptide-membrane interface. The peptide was considered to have eliminated the water layer once the solvation index reached a stable value. In all five runs, there is a degree of oscillation in the solvation index, making the discernment of the precise point of desolvation difficult. The point of desolvation is further obscured by fact that the peptide may also be reorienting to where residues with differing hydrophilicity may face the peptide-membrane interface. As desolvation of the water layer should occur from the cessation of the scaling of peptide-environment interactions, estimates should only be taken from simulations that show distinct drops and stabilization of their solvation index around the beginning of the simulations. This will, however, cause the estimates of desolvation time to be lower than what they might actually be. Simulations 3 and 4 show a clear trend in their

solvation and, from these two simulations, the desolvation time of Gm was estimated to be around 20 ns.

The EAF needed to allow a structure to transition from the highest to the ground replica to achieve desolvation was calculated with the following formula:

$$EAF = \frac{t*P(a)}{R} \quad \text{(Eq. 9)}$$

where t is the time taken (in this case, the desolvation time) by the configuration to transition from the highest to the ground replica, P(a) is the acceptance probability, and R is the number of replicas. This assumes that in each replica of the REST3 regime, the peptide-membrane interface has the same tendency to desolvate as in an unscaled simulation. As the magnitude of the peptide-environment interactions are reduced in all replicas apart from the ground replica, this assumption does not hold and the EAF calculated will be an underestimate. Given a desolvation time of 20 ns, 10 replicas and an average acceptance probability of 20%, the minimum required EAF would be 400 ps. With this EAF, across a 50-ns replica exchange simulation there will be only 125 attempted exchanges. Given an acceptance probability of 0.2, only 25 exchanges would be made successfully. Considering the computational resources available to us, this is beyond what would be deemed efficient to use across the entire reaction coordinate for the calculation of the PMF.

To confirm that an estimated minimum value of 400 ps is not in fact faster, a replica exchange regime identical to the previously used Max SF 0.2 was conducted, except that an EAF of 500 ps was used. An EAF of 500 ps was used instead of 400 ps to compensate for the value of 400 ps likely being an underestimate of the required minimum time for desolvation. The peptide was held at a COM-distance of 2.6 nm. Figure 8B shows histograms of the solvation index for the ground replica of the replica exchange regimes with a Max SF of 0.2, and with an EAF of 0.1 ps or 500 ps, and the equivalent number of unscaled replicate simulations as described in section 3.4. The unscaled simulations were computed to have a median solvation index of around 0.19 ± 0.04 water molecules per A^2. The regime with an EAF of 0.1 ps has a median of 0.48 ± 0.11 water molecules per A^2. This density of interfacial water molecules is sufficient to cause repulsion in the PMF (Fig. 8). Increasing the EAF to 500 ps caused only a relatively modest reduction in the solvation index, with a median of 0.39 ± 0.12 waters per A^2. This indicates that an EAF of 500 ps (or lower) is not sufficient to allow for the peptide to eliminate the solvation layer. Given that this EAF is not efficient enough

for windows across the entire reaction coordinate, longer EAFs were not investigated. Therefore, increasing the EAF does not appear to be a feasible approach to address the problem of the formation of a solvation layer within the REST3 scheme.

3.7 Peptide-peptide interaction scaling

As previously shown, when the peptide-environment interactions are scaled, the peptide compensates for the loss of external interactions by gaining internal HBs (Fig. 6F). It was hypothesised that the increase of peptide-membrane HBs in the ground replica of the Max SF 0.2 regime was reduced by the increase in intra-peptide HBs. Therefore, by reducing the peptide-peptide interactions whilst scaling the peptide-environment interactions as before, the peptide may return to its native conformation and Gm would be better able to regain HBs with the membrane, as a limited reduction in internal Gm-Gm interactions would facilitate Gm-membrane HBs to be regained.

Before incorporating peptide-peptide scaling into the computationally-expensive replica exchange regime, scaled simulations of the peptide held at a COM-distance of 2.6 nm (membrane surface) were conducted for 200 ns with peptide-peptide interaction SFs of 0.8, 0.6, and 0.4 (Table 1: sPP-0.8 → sPP-0.4). In all of these simulations, the peptide-environment interactions were scaled by 0.4 to understand the effect of the magnitude of the peptide-peptide scaling on the structure of the peptide.

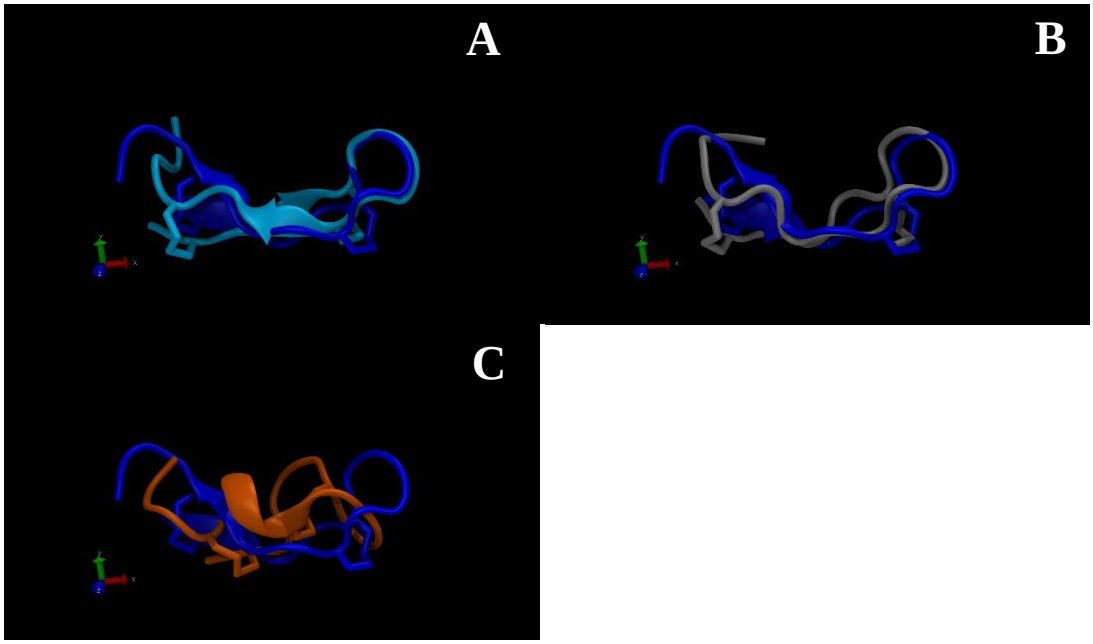


Figure 11: Comparison of the structure of Gm in simulations with peptide-environment interactions that are scaled to 0.4 (blue) and with scaling peptide-peptide interaction SFs of A: 0.8, B: 0.6, C: 0.4 (Table 1: sPP-0.8 → sPP-0.4). In all simulations the peptide was held at a COM distance of 2.6 nm to the membrane. The membrane is not shown in these images for clarity. The structures presented here are not necessarily representative of the conformations of Gm across the simulation trajectory, and are only examples of the deformations that can occur.

When the peptide-environment interactions were scaled by 0.4 and the peptide-peptide interactions were left unscaled, the peptide exhibited an increased median number of 16.3 ± 2.3 Gm-Gm HBs, compared to 7.0 ± 1.4 in unscaled simulations. It should be reiterated that the change in structure associated with the increased number of HBs is not desirable because Gm has a rigid β-hairpin structure that is known to persist independently of its environment.(21, 37, 62) When the peptide-peptide interactions were reduced with a SF of 0.8, Gm-Gm interactions stayed relatively the same at a median of 15.0 ± 2.3 HBs. Reductions in the median number of Gm-Gm HBs to 12.5 ± 2.5 and 8.6 ± 2.2 were observed when the SF was reduced to 0.6 and 0.4, respectively. Whilst scaling the peptide-peptide interactions appears to counteract the gain in Gm-Gm HBs due to the scaling of peptide-environment interactions, this does not, however, also counteract the unwanted changes in the structure of the peptide. It was observed that there are only minimal changes to the structure of the peptide in simulations with peptide-peptide interaction SFs of 0.8 and 0.6 compared to a simulation where only peptide-environment interactions are scaled down (Fig. 11A-B). Dramatic changes, such as the formation of an α-helix in the structure of Gm, were observed in the simulation with a peptide-peptide interaction SF of 0.4 (Fig. 11C). It is likely that the simple scaling down of the peptide-peptide interactions to counteract the increased Gm-Gm HBs is not sufficient to see a return of the native structure of Gm. This is because the structure of the peptide is not solely dependent on the internal peptide interactions, but also the interactions of the peptide with water. As the peptide-water interactions remain scaled, a hydrophobic effect will be created between the peptide and water molecules, contributing to the disturbance in the structure of Gm. Thus, scaling the peptide-peptide interactions does not counteract the structural changes caused by the scaling of the peptide-environment interactions, resulting in either the same or further structural disturbances. Since this would neither aid in increasing the rate at which the water layer at the peptide-membrane interface is eliminated nor result in a more accurate structure of the peptide, this approach was not considered suitable.

3.8 Membrane-water interaction scaling

Membrane-water scaling was conducted to test if this would reduce the formation of the water layer whilst also allowing scaling of the peptide-environment interactions to permit configurational sampling. Membrane-water interaction SFs of 0.8, 0.6, 0.4, and

0.2 were tested at a COM-distance of 2.6 nm in simulations each of 200 ns (Table 1: sMW-0.8 → sMW-0.2). The peptide-environment interactions were left unscaled.

As before, analysis of HBs was used to assess the effect of the scaling. It was observed that scaling of the membrane-water interactions led to the peptide rapidly gaining HBs with the membrane (Fig. S7A in the Supporting Information). Given that it is likely that the loss of peptide-membrane interactions that enables Gm to sample new configurations whilst scaling of peptide-environment interactions (Fig. 6), if membrane-water and peptide-environment interaction scalings were used concurrently, the peptide would either sample new configurations or remain bound to the membrane. The behaviour of the peptide will depend on the magnitude of the SFs, and which ever effect dominates will defeat the purpose of using the opposing scaling condition.

Scaling of membrane-water interactions was also observed to be accompanied by a compression of the membrane and a decrease in the area per lipid (APL). The median APL in the unscaled simulation was 0.59 nm^2, which decreases with the membrane-water SF to 0.52 ± 0.01, 0.48 ± 0.01, 0.47 ± 0.01, and 0.45 ± 0.01 nm^2 for membrane-water interaction SFs of 0.8, 0.6, 0.4, and 0.2, respectively. This effect is undesirable, as a correct membrane structure is necessary for an accurate computation of peptide-membrane interactions. APL is often used to track the equilibration of a membrane during a MD simulation. These equilibration periods can often last hundreds of nanoseconds, and if a replica exchange regime were to feature a replica with membrane-water scaling that significantly modifies the APL, the membrane will likely not return to its equilibrium state by the time the configuration reaches the ground replica. In addition to the reduction in the APL, an expected accompanying increase in the *z* dimension of the simulation box was observed. For each of the simulations with membrane-water interaction SFs starting and descending from 1.0, the median *z* length of the simulation cell was 14.7 ± 0.2, 16.9 ± 0.2, 18.1 ± 0.2, 18.5 ± 0.1, and 19.8 ± 0.3 nm, respectively. As scaling the membrane-water interactions caused changes to the APL and the simulation cell dimensions, it was concluded that it is not appropriate to scale interactions in this manner.

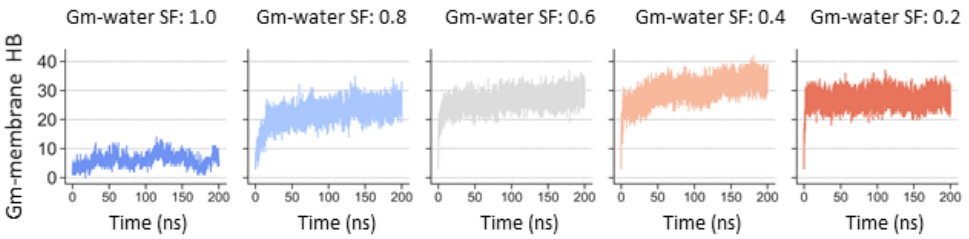
3.9 Effect of unequal peptide-membrane and peptide-water scaling

Scaling Peptide-peptide and membrane-water interactions was not effective at eliminating the water layer at the peptide-membrane interface without producing deformations of the structure of the peptide or the membrane. Alternative approaches

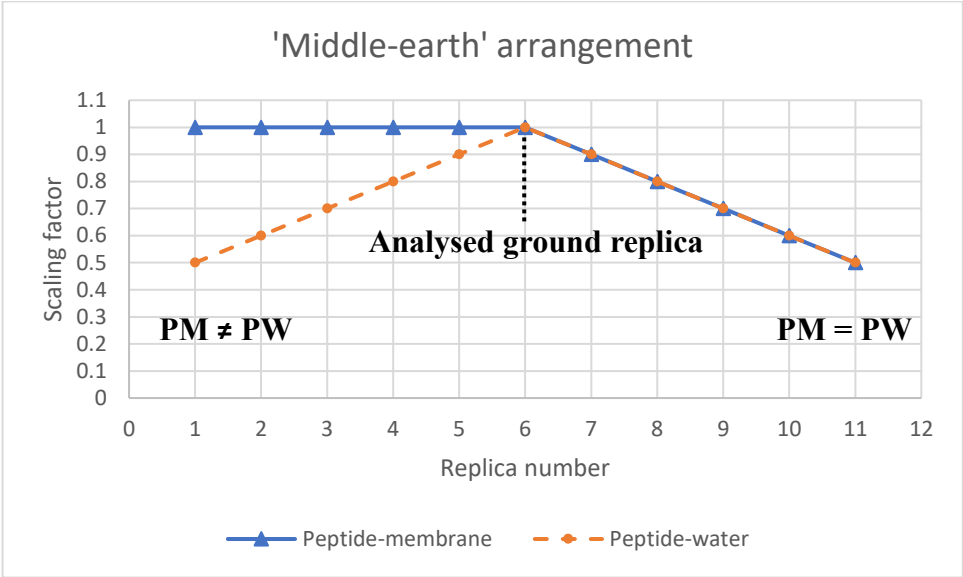
that involve different scalings of the peptide-membrane and peptide-water SF would be needed. As this would conflict with an equal reduction of peptide-environment interactions needed to sample new configurations, such an alternative approach would need to be included into a REST3 simulation regime in separate replicas. The approach investigated here was to scale the peptide-membrane and peptide-water interactions such that the SF of the peptide-membrane interaction was greater than the SF of the peptide-water interaction. This type of peptide-environment scaling is referred to as 'decoupled' SF. The approach described so far, where the SF was the same for both peptide-membrane and peptide-water interactions is referred to as 'equal' SF. The reasoning behind the decoupled approach is that by decoupling the SF of peptide-membrane and peptide-water interactions, the balance of interactions might shift in favour of the peptide gaining interactions with the membrane rather than with water.

To test this approach, five decoupled-scaling simulations were performed (with no replica exchange) (Table 1: simulations sUe-0.8 → sUe-0.2). In all of them, the peptide-membrane interaction SF was left unscaled (SF = 1.0) whilst the peptide-water interaction SFs were 1.0, 0.8, 0.6, 0.4 or 0.2. (Fig. 12A). The peptide was held at a COM-distance of 2.6 nm and each simulation was run for 200 ns. Figure 12A shows the median number of Gm-membrane HBs across time for these decoupled scaling simulations. The data demonstrates that decoupling the peptide-environment SFs rapidly increases the number of HBs between the peptide and membrane. In the unscaled simulation, the number of Gm-membrane HBs remain mostly below 10 HBs with a median of 6.0 ± 2.1 HBs. In the decoupled simulations, the median numbers of Gm-membrane HBs are 23.0 ± 4.0, 27.0 ± 3.0, 31.0 ± 3.9, and 27.0 ± 2.5, for peptide-water interaction SFs of 0.8, 0.6, 0.4, and 0.2, respectively. The time it takes for Gm-membrane HBs to form also decreases as the peptide-water SF decreases. It takes less than 5 ns for the number of Gm-membrane HBs to stabilise in decoupled simulations with peptide-water interaction SFs of 0.6, 0.4 and 0.2. This increases to 25 ns for the decoupled simulation with a peptide-water interaction SF of 0.8 (although there is still a slight upward drift over the remainder of the simulation). Fig. S8 in the Supporting Information contains plots showing the changes in the number of HBs between the other system components.

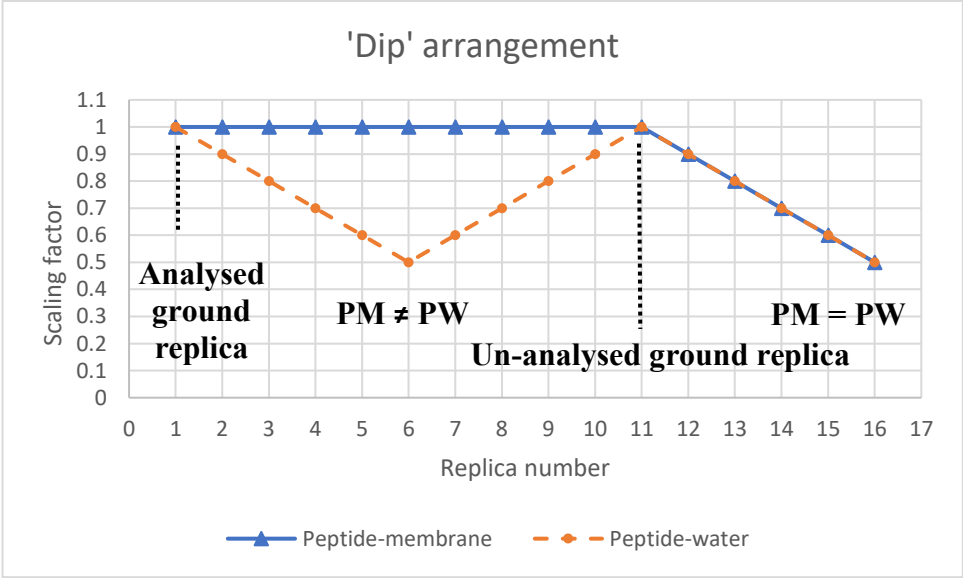
A

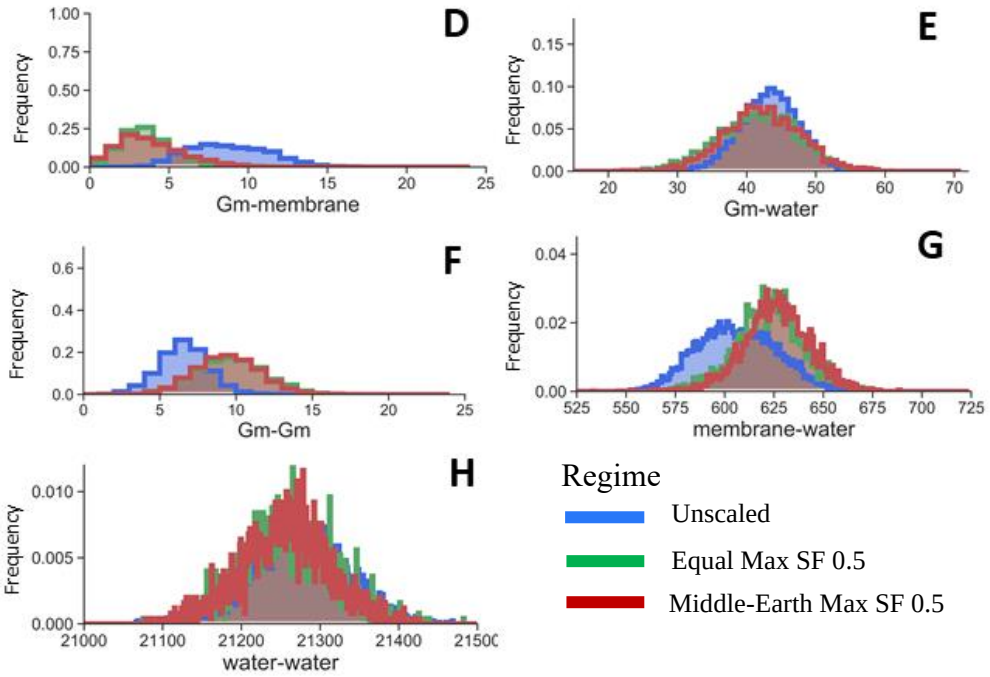

Gm-water SF: 1.0
Gm-water SF: 0.8
Gm-water SF: 0.6
Gm-water SF: 0.4
Gm-water SF: 0.2
Gm-membrane HB
Time (ns)

B


'Middle-earth' arrangement
Scaling factor
Analysed ground replica
PM ≠ PW
PM = PW
Replica number
Peptide-membrane
Peptide-water

C


'Dip' arrangement
Scaling factor
Analysed ground replica
PM ≠ PW
PM = PW
Un-analysed ground replica
Replica number
Peptide-membrane
Peptide-water



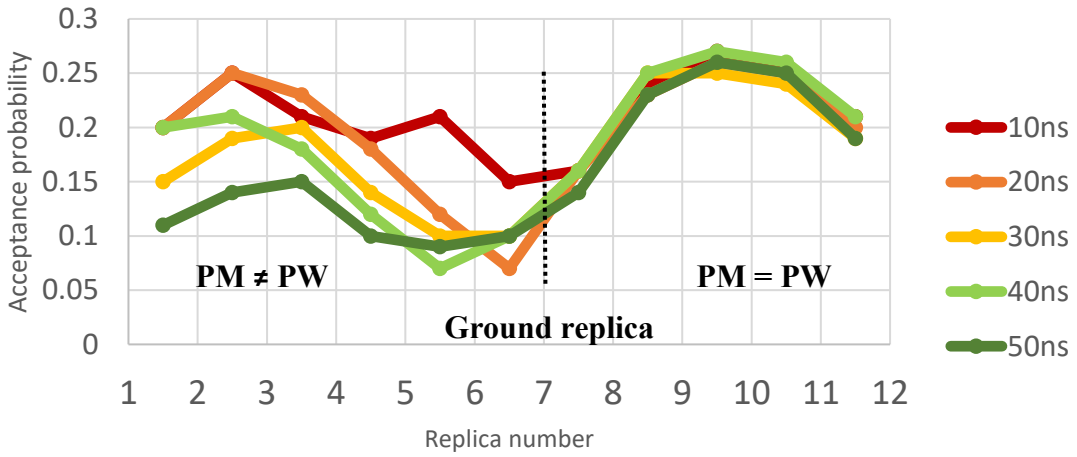


Figure 12: A: Evolution of the average number of Gm-membrane hydrogen bonds across time for peptide-membrane (PM) interaction SF = 1.0 and when the peptide-water (PW) interaction SF is reduced to 0.8, 0.6, 0.4, and 0.2 (Table 1: sUE-0.8 → sUE-0.2). In these simulations the peptide is held at a COM distance of 2.6 nm to the membrane. B: Hypothetical REST3 regime showing an arrangement of scaling factors termed the 'middle-earth' arrangement, that possesses a maximum decoupled SF and equal SF of 0.5. The blue and orange lines show how the PM and PW interaction SFs change across replicas, respectively. C: Hypothetical REST3 regime in a 'dip' arrangement. Normalised histograms of the number of D: Gm-membrane, E: Gm-water, F: Gm-Gm, G: membrane-water, and H: water-water HBs. In blue, is the HB from an unscaled simulation (Table 1: cUS-2.6). In green is a REST3 simulation with a maximum SF of 0.5, where peptide-membrane and peptide-water interactions are scaled equally throughout. In red is a REST3 regime arranged in a middle earth arrangement (Table 1: rME-0.5). The unequal peptide-membrane and peptide-water SF side and the equal side both go to maximum SFs of 0.5. I: Average exchange acceptance probability (ratio) across replicas of a middle-earth arranged regime with a max SF of 0.5 (Table 1: rME-0.5). Each

colour represents the average over the corresponding time segment of the simulation as described in the figure legend.

To incorporate the above decoupled scaling approach into a REST3 replica exchange regime, it must be decided what set of decoupled peptide-environment interaction SFs to use, such that the configurations in the ground replica do not feature a water layer or overestimated Gm-membrane interactions. Ideally, to determine if the choice in decoupled SFs correctly counteracts the loss of interactions caused by equal SFs, a full REST3 simulation would be conducted for a potential set of decoupled SFs to obtain a full PMF. This would allow a comparison to experiment of the ΔG_b predicted using different decoupled SFs and an assessment of the accuracy of the interactions in the system. This is, however, computationally prohibitive. It is more feasible to use some property (or properties) of the system to assess the effect of a set of chosen decoupled SFs on the interactions in the system before undertaking a full REST3 calculation, so that it could be reasonably assessed that an accurate PMF may be calculated. Given that the main aim of the decoupled scaling approach is to counteract the formation of the extraneous water layer and the loss of Gm-membrane interactions, potential properties to be considered include the solvation index and/or the number of Gm-membrane HBs. However, the problem with this approach is that knowledge is first needed of what constitutes a 'reasonable' number of peptide-membrane interfacial water molecules or Gm-membrane HBs, for any particular configuration. In other words, the 'target' number of interactions to be achieved by the use of decoupled SFs needs to be established first. This *a priori* knowledge does not exist, and so it is not possible to decide upon an appropriate set of decoupled SFs to use without the prohibitive calculation of the full PMF for a large set of possible decoupled SFs.

Even if this were computationally feasible, there still lies the question of how a selected set of SFs would be implemented within a replica exchange regime. This itself is not necessarily prohibitively expensive to test. There are two possible approaches: the so-called 'middle-earth' and the 'dip' arrangements. In the middle-earth arrangement, the ground (or 'earth') replica is located between a set of decoupled SFs on one side and a set of equal SFs on the other, with the magnitude of the scaling increasing the further the replica is from the ground replica (Fig. 12B).(84) In the 'dip' arrangement, there are two ground replicas, although only one is analysed (Fig. 12C). Between the two ground replicas the peptide-water interaction SF 'dips', such that as

the SF decreases from the first ground replica until it reaches the maximum decoupled SF, and then increases as it approaches the next ground replica. The peptide-membrane interaction SF remains at a value of 1.0 across the 'dip', creating a region of decoupled scaling. From one of the ground replicas, a set of equal SFs is connected, where the peptide-environment interaction SFs decrease the further the replica is from the connected ground replica until the maximum equal SF is reached. The analysed ground replica is the one that is not directly connected to the equal SF set.

For the same set of maximum SFs, the 'middle-earth' arrangement has fewer replicas than the 'dip' arrangement, thus using fewer computational resources. However, in the 'middle-earth' arrangement, the configurations that are exchanged into the ground replica may come from the equal SF side or the decoupled SF side, without the configurations from the equal SF side necessarily passing through the decoupled SF, or *vice versa*, before exchanging into the ground replica. Thus, the ground replica will feature configurations that have under-estimated peptide-membrane interactions, over-estimated peptide-membrane interactions, as well as 'corrected' ones. This issue is partially mitigated by the 'dip' arrangement, as configurations coming from the equal SF side must first pass through the decoupled SF replicas before they can reach the analysed ground replica. The 'dip' arrangement will, however, still receive configurations that have only traversed through the decoupled SF replicas and thus will also feature over-estimated, favourable peptide-membrane interactions. Therefore, it was hypothesised that decoupled SFs cannot be implemented into a REST3 regime without over-estimating favourable peptide-membrane interactions. This problem would apply similarly to any type of SFs introduced into a regime to counteract the behaviour of another group of replicas within the regime.

To investigate how the arrangement of decoupled SFs would affect the outcome of the replica exchange, a simulation with a 'middle-earth' arrangement was conducted (Table 1: simulation rME-0.5). On the side of equal peptide-environment interaction SF replicas, the SF was decreased in increments of 0.1 to a maximum value of 0.5. The decoupled peptide-environment interaction set of SFs was scaled such that the peptide-membrane interaction SFs were kept at 1.0, as the peptide-water interaction SFs were reduced to 0.5 in increments of 0.1. The maximum SFs were chosen arbitrarily, as it is not possible to determine the necessary values of decoupled SFs, as argued previously. A HB analysis of this simulation with a 'middle-earth' arrangement showed no difference in any of the numbers of HBs between system components to

the numbers of HBs seen in an analogous regime that featured only equal peptide-environment interaction SFs to a Max SF of 0.5 (Fig. 12D-H). This appears to be due to poor exchange rates observed between decoupled SF replicas in the 'middle-earth' arrangement regime. Figure 12I shows plots of the exchange acceptance probability (ratio) vs replica number for sequential time 'slices' across the trajectory. The exchange acceptance probability across replicas is broken up into 10-ns slices. The point between two replica numbers is the probability of an exchange being accepted between said replicas over the 10-ns slice. During the first 10 ns, exchange acceptance ratios were initially at an acceptable level of around 0.15 to 0.25. As the simulation progressed, the ratios on the decoupled SF side reduced and fluctuated to a larger degree relative to the equal SF replica set. By the last slice of the simulations, the ratios of the decoupled SF replica set were between 0.10 and 0.15. Due to this drop in exchange acceptance ratios on the decoupled SF side of the regime and the relatively higher exchange acceptance ratio between the equal SF replicas, the configurations exchanged into the ground replica were found to come overwhelmingly from the equal SF replicas. As a result, no change was observed in the pattern of HBs between the 'middle-earth' arrangement regime and its analogous equal SF only regime. To mitigate such an issue, given the reduction in exchange acceptance ratio to around 0.1 – 0.15, an exchange acceptance ratio of around 0.3-0.35 should be aimed for when attempting to construct a decoupled scaling REST3 regime. To do so, however, would require a larger number of replicas to be included, increasing the computational cost. Thus, it was deemed not economical, nor fruitful given the theoretical problems with assembling a REST3 regime featuring decoupling, to continue further testing.

4.0 Conclusions

To characterise the interactions that govern membrane binding of Gm and thus assist the rational design of peptides with specific properties, simulations to predict ΔG_b must first be validated by comparing predicted binding affinities to experimental data. The PMF of the binding of Gm to a POPC bilayer was first calculated using conventional MD simulations with US, and it was found that the system does not sample peptide-membrane configurations sufficiently enough to give an accurate estimation of ΔG_b. The estimated value of ΔG_b was -60 ± 5 kJ mol^{-1}, which is a severe overestimation of the binding affinity for a peptide that is likely to only show weak binding to a POPC

bilayer.(43) Therefore, a newly developed enhanced sampling technique called REST3 was used, which allows the selective scaling of interactions between any two system components.

REST3 was found to be able to increase the sampling of peptide-membrane configurations, as desired; however, the PMF computed from the use of REST3-US simulations suggested that Gm is repelled by POPC bilayers. This is not in accordance with experimental data suggesting Gm to be a weak binder to POPC membranes, nor with the behaviour of the peptide in conventional, unscaled MD simulations.(78) The lack of binding was determined to be caused by the scaling of peptide-environment interactions. When the peptide-membrane and peptide-water interactions are scaled equally, the membrane has more favourable interactions with water molecules than it does with the peptide. As a result, the water molecules displace the peptide from the membrane surface and a layer of interfacial water between the membrane and the peptide forms. This water layer blocks the formation of favourable peptide-membrane interactions. When these configurations are exchanged down into the ground replica, there is insufficient time for the peptide to displace these extraneous water molecules. Hence, the PMF computed exhibits a strong repulsion between Gm and the membrane. A number of approaches were investigated to try to solve this problem.

Increasing the amount of time between attempted replica exchanges was initially investigated as an approach to allow time for the desolvation of the peptide-membrane interface. It was estimated that Gm takes approximately 20 ns to displace the solvation layer. To allow for this period of time to occur in a replica exchange regime, an exchange attempt frequency of over 0.5 ns would be needed. This would require each window of a REST3-US simulation to be run for tens of microseconds to obtain a reasonable amount of replica exchanges, making the approach inefficient and unfeasible with current computational resources. Alternative scaling methods were then investigated in an attempt to reduce the desolvation time. Scaling the peptide-peptide or membrane-water interactions whilst also reducing the peptide-environment interactions was found to either not improve or to introduce further structural deformations to the peptide or membrane. Thus, it was concluded that hindering the formation of the water layer could not be achieved whilst simultaneously scaling the peptide-water and peptide- membrane interactions to equal levels.

A decoupled approach was subsequently tested whereby the peptide-membrane interaction SF was chosen to be greater than the peptide-water interaction SF. It was

found that this approach caused Gm to rapidly gain HBs with the membrane, presumably removing the solvation layer as well. However, the implementation of such decoupled SFs proved to be difficult. The selection of an appropriate SF cannot be done without the calculation of a full PMF, which is not feasible for all the potential sets of decoupled SFs that should be tested. It was also hypothesised that the inclusion of decoupled SFs into a REST3 regime would necessarily result in configurations with over-estimated peptide-membrane interactions featuring in the ground replica. This could not be confirmed as the exchange acceptance ratios between the decoupled replicas was significantly reduced over the course of the test simulations. This suggested that higher exchange acceptance ratios than what is generally considered efficient would need to be used to optimise such a decoupled replica exchange regime. The increased number of replicas needed to accommodate the increased exchange acceptance ratio would also further reduce the efficiency of the method. It was thus concluded that continued testing of this approach is not worthwhile.

The slow rate of desolvation of the peptide may be mitigated with the use of a different force field. A recently published modification to the PC lipid headgroup for GROMOS compatible parameters may offer such a possibility.(82) This parameter set reduces the magnitude of the headgroup charges and more accurately describes the interaction of a peptide with a PC membrane. Due to the reduced interactions with the membrane and, therefore, the likely presence of more water molecules at the interface in this force field, its use might mitigate the overestimation of the binding of the peptide in a conventional US calculation of the PMF. It may also reduce the time taken for the peptide-membrane interface to desolvate by reducing the number of interfacial water molecules that must be removed before reaching a stable solvation index. Hence, a faster EAF may be used (according to eq. 9) and use of a REST3 regime might be within a more permissible range of efficiency.

While REST3 can enhance the sampling of the Gm peptide, the presence of the water layer prevented an accurate calculation of ΔG_b. Efforts were directed towards preventing the formation of this water layer and, as this was not successful, simulations of furthering membrane complexity or different Gm variants were not conducted. As such, the simulation could not be validated as was originally intended, and the molecular mechanism of binding of Gm could not be elucidated.

Currently the REST3 method does not seem suitable for use with the Gomesin-POPC system. The issues with, and trying to solve, the loss of peptide-membrane interactions

and the associated formation of a water layer would likely affect any other multi-component systems, where there are numerous and slow-equilibrating interactions between the system components. Increased computational power may help to alleviate the slow desolvation rate by allowing slower EAFs, but currently the available resources restrict further investigation. Whilst it is not recommended that REST3 be used for other ACPs, AMPs or similar systems, the method may still be suitable for use with small molecules or peptides of only a few residues.

5.0 References

1. Felício MR, Silva ON, Gonçalves S, Santos NC, Franco OL. Peptides With Dual Antimicrobial And Anticancer Activities. Frontiers in Chemistry. 2017;5(5).
2. Gaspar D, Veiga AS, Castanho MARB. From Antimicrobial To Anticancer Peptides. A Review. Frontiers in Microbiology. 2013;4:294.
3. Schweizer F. Cationic Amphiphilic Peptides With Cancer-Selective Toxicity. European Journal of Pharmacology. 2009;625(1):190-4.
4. Riedl S, Zweytick D, Lohner K. Membrane-Active Host Defense Peptides – Challenges And Perspectives For The Development Of Novel Anticancer Drugs. Chemistry and Physics of Lipids. 2011;164(8):766-81.
5. Torres MDT, Sothiselvam S, Lu TK, de la Fuente-Nunez C. Peptide Design Principles For Antimicrobial Applications. Journal of Molecular Biology. 2019;431(18):3547-67.
6. Van Gunsteren WF, Bakowies D, Baron R, Chandrasekhar I, Christen M, Daura X, et al. Biomolecular Modeling: Goals, Problems, Perspectives. Angewandte Chemie International Edition. 2006;45(25):4064-92.
7. Nowak W. Applications Of Computational Methods To Simulations Of Proteins Dynamics. In: Leszczynski J, Kaczmarek-Kedziera A, Puzyn T, G. Papadopoulos M, Reis H, K. Shukla M, editors. Handbook of Computational Chemistry. Cham: Springer International Publishing; 2017. p. 1627-69.
8. Van Gunsteren WF, Daura X, Hansen N, Mark AE, Oostenbrink C, Riniker S, et al. Validation Of Molecular Simulation: An Overview Of Issues. Angewandte Chemie International Edition. 2018;57(4):884-902.
9. Kästner J. Umbrella Sampling. Wiley Interdisciplinary Reviews: Computational Molecular Science. 2011;1(6):932-42.
10. Neale C, Pomès R. Sampling errors in free energy simulations of small molecules in lipid bilayers. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2016;1858(10):2539-48.
11. Martinotti C, Ruiz-Perez L, Deplazes E, Mancera RL. Molecular Dynamics Simulation of Small Molecules Interacting with Biological Membranes. ChemPhysChem. 2020;21(14):1486-514.
12. Bussi G, Laio A. Using metadynamics to explore complex free-energy landscapes. Nature Reviews Physics. 2020;2(4):200-12.
13. Pokhrel N, Maibaum L. Free Energy Calculations of Membrane Permeation: Challenges Due to Strong Headgroup–Solute Interactions. Journal of Chemical Theory and Computation. 2018;14(3):1762-71.
14. Nguyen LT, Haney EF, Vogel HJ. The Expanding Scope Of Antimicrobial Peptide Structures And Their Modes Of Action. Trends in Biotechnology. 2011;29(9):464-72.
15. Cao Z, Liu L, Hu G, Bian Y, Li H, Wang J, et al. Interplay of hydrophobic and hydrophilic interactions in sequence-dependent cell penetration of spontaneous membrane-translocating peptides revealed by bias-exchange metadynamics simulations. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2020;1862(10):183402.
16. Piana S, Laio A. A Bias-Exchange Approach to Protein Folding. The Journal of Physical Chemistry B. 2007;111(17):4553-9.
17. Nielsen SO, Lopez CF, Srinivas G, Klein ML. Coarse grain models and the computer simulation of soft materials. Journal of Physics: Condensed Matter. 2004;16(15):R481-R512.
18. Bond PJ, Holyoake J, Ivetac A, Khalid S, Sansom MSP. Coarse-grained molecular dynamics simulations of membrane proteins and peptides. Journal of Structural Biology. 2007;157(3):593-605.
19. Arnarez C, Uusitalo JJ, Masman MF, Ingólfsson HI, de Jong DH, Melo MN, et al. Dry Martini, a Coarse-Grained Force Field for Lipid Membrane Simulations with Implicit Solvent. Journal of Chemical Theory and Computation. 2015;11(1):260-75.

20. Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink S-J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. Journal of Chemical Theory and Computation. 2008;4(5):819-34.
21. Moraes LG, Fazio MA, Vieira RF, Nakaie CR, Miranda MT, Schreier S, et al. Conformational And Functional Studies Of Gomesin Analogues By Cd, Epr And Fluorescence Spectroscopies. Biochimica et Biophysica Acta. 2007;1768(1):52-8.
22. Huang Y, Huang J, Chen Y. Alpha-helical cationic antimicrobial peptides: relationships of structure and function. Protein & Cell. 2010;1(2):143-52.
23. Houston Jr ME, Kondejewski LH, Karunaratne DN, Gough M, Fidai S, Hodges RS, et al. Influence of preformed α-helix and α-helix induction on the activity of cationic antimicrobial peptides. The Journal of Peptide Research. 1998;52(2):81-8.
24. Sugita Y, Okamoto Y. Replica-Exchange Molecular Dynamics Method For Protein Folding. Chemical Physics Letters. 1999;314(1):141-51.
25. Liu P, Kim B, Friesner RA, Berne BJ. Replica Exchange With Solute Tempering: A Method For Sampling Biological Systems In Explicit Water. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(39):13749-54.
26. Wang L, Friesner RA, Berne BJ. Replica Exchange With Solute Scaling: A More Efficient Version Of Replica Exchange With Solute Tempering (Rest2). Journal of Physical Chemistry B. 2011;115(30):9431-8.
27. Olson MA, Lee MS, Yeh I-C. Membrane insertion of fusion peptides from Ebola and Marburg viruses studied by replica-exchange molecular dynamics simulations. Journal of Computational Chemistry. 2017;38(16):1342-52.
28. Baldissera G, dos Santos Cabrera MP, Chahine J, Ruggiero JR. Role of Peptide–Peptide Interactions in Aggregation: Protonectins Observed in Equilibrium and Replica Exchange Molecular Dynamics Simulations. Biochemistry. 2015;54(13):2262-9.
29. Im W, Brooks CL. Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(19):6771-6.
30. Im W, Feig M, Brooks CL, 3rd. An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. Biophysical Journal. 2003;85(5):2900-18.
31. Jaskierny AJ, Panahi A, Feig M. Effect of flanking residues on the conformational sampling of the internal fusion peptide from Ebola virus. Proteins: Structure, Function, and Bioinformatics. 2011;79(4):1109-17.
32. Guo C, Côté S, Mousseau N, Wei G. Distinct Helix Propensities and Membrane Interactions of Human and Rat IAPP1–19 Monomers in Anionic Lipid Bilayers. The Journal of Physical Chemistry B. 2015;119(8):3366-76.
33. Côté S, Binette V, Salnikov Evgeniy S, Bechinger B, Mousseau N. Probing the Huntingtin 1-17 Membrane Anchor on a Phospholipid Bilayer by Using All-Atom Simulations. Biophysical Journal. 2015;108(5):1187-98.
34. Smith AK, Klimov DK. Binding of Cytotoxic Aβ25–35 Peptide to the Dimyristoylphosphatidylcholine Lipid Bilayer. Journal of Chemical Information and Modeling. 2018;58(5):1053-65.
35. Smith AK, Klimov DK. De novo aggregation of Alzheimer's Aβ25-35 peptides in a lipid bilayer. Scientific Reports. 2019;9(1):7161.
36. Silva PI, Jr., Daffre S, Bulet P. Isolation And Characterization Of Gomesin, An 18-Residue Cysteine-Rich Defense Peptide From The Spider Acanthoscurria Gomesiana Hemocytes With Sequence Similarities To Horseshoe Crab Antimicrobial Peptides Of The Tachyplesin Family. Journal of Biological Chemistry. 2000;275(43):33464-70.
37. Mandard N, Bulet P, Caille A, Daffre S, Vovelle F. The Solution Structure Of Gomesin, An Antimicrobial Cysteine-Rich Peptide From The Spider. European Journal of Biochemistry. 2002;269(4):1190-8.

38. Panteleev PV, Bolosov IA, Balandin SV, Ovchinnikova TV. Structure And Biological Functions Of B-Hairpin Antimicrobial Peptides. Acta Naturae. 2015;7(1):37-47.
39. Buri MV, Dias CC, Barbosa CM, Nogueira-Pedro A, Ribeiro-Filho AC, Miranda A, et al. Gomesin Acts In The Immune System And Promotes Myeloid Differentiation And Monocyte/Macrophage Activation In Mouse. Peptides. 2016;85:41-5.
40. Lorenzini DM, Fukuzawa AH, da Silva PI, Machado-Santelli G, Bijovsky AT, Daffre S. Molecular Cloning, Expression Analysis And Cellular Localization Of Gomesin, An Anti-Microbial Peptide From Hemocytes Of The Spider Acanthoscurria Gomesiana. Insect Biochemistry and Molecular Biology. 2003;33(10):1011-6.
41. World_Health_Organisation. Global Priority List Of Antibiotic-Resistant Bacteria To Guide Reasearch, Discovery, And Development Of New Antibiotics. Geneva; 2017.
42. Rodrigues EG, Dobroff AS, Cavarsan CF, Paschoalin T, Nimrichter L, Mortara RA, et al. Effective Topical Treatment Of Subcutaneous Murine B16f10-Nex2 Melanoma By The Antimicrobial Peptide Gomesin. Neoplasia. 2008;10(1):61-8.
43. Henriques TS, Lawrence N, Chaousis S, Ravipati AS, Cheneval O, Benfield AH, et al. Redesigned Spider Peptide with Improved Antimicrobial and Anticancer Properties. ACS Chemical Biology. 2017;12(9):2324-34.
44. Chan LY, Zhang VM, Huang YH, Waters NC, Bansal PS, Craik DJ, et al. Cyclization Of The Antimicrobial Peptide Gomesin With Native Chemical Ligation: Influences On Stability And Bioactivity. ChemBioChem. 2013;14(5):617-24.
45. Ikonomopoulou MP, Fernandez-Rojo MA, Pineda SS, Cabezas-Sainz P, Winnen B, Morales RAV, et al. Gomesin Inhibits Melanoma Growth By Manipulating Key Signaling Cascades That Control Cell Death And Proliferation. Scientific Reports. 2018;8(1):11519.
46. Buri MV, Torquato HFV, Barros CC, Ide JS, Miranda A, Paredes-Gamero EJ. Comparison Of Cytotoxic Activity In Leukemic Lineages Reveals Important Features Of Beta-Hairpin Antimicrobial Peptides. Journal of Cellular Biochemistry. 2017;118(7):1764-73.
47. Soletti RC, del Barrio L, Daffre S, Miranda A, Borges HL, Moura-Neto V, et al. Peptide Gomesin Triggers Cell Death Through L-Type Channel Calcium Influx, Mapk/Erk, Pkc And Pi3k Signaling And Generation Of Reactive Oxygen Species. Chemico-Biological Interactions. 2010;186(2):135-43.
48. Schweizer F. Cationic amphiphilic peptides with cancer-selective toxicity. European Journal of Pharmacology. 2009;625(1-3):190-4.
49. Hoskin DW, Ramamoorthy A. Studies On Anticancer Activities Of Antimicrobial Peptides. Biochimica et Biophysica Acta. 2008;1778(2):357-75.
50. Tanner J, Deplazes E, Mancera R. The Biological And Biophysical Properties Of The Spider Peptide Gomesin. Molecules. 2018;23(7):1733.
51. Gaspar D, Veiga AS, Castanho MARB. From antimicrobial to anticancer peptides. A review. Frontiers in Microbiology. 2013;4.
52. Harris F, Dennison SR, Singh J, Phoenix DA. On The Selectivity And Efficacy Of Defense Peptides With Respect To Cancer Cells. Medicinal Research Reviews. 2013;33(1):190-234.
53. Mattei B, Miranda A, Perez KR, Riske KA. Structure-Activity Relationship Of The Antimicrobial Peptide Gomesin: The Role Of Peptide Hydrophobicity In Its Interaction With Model Membranes. Langmuir. 2014;30(12):3513-21.
54. Domingues TM, Mattei B, Seelig J, Perez KR, Miranda A, Riske KA. Interaction Of The Antimicrobial Peptide Gomesin With Model Membranes: A Calorimetric Study. Langmuir. 2013;29(27):8609-18.
55. Domingues TM, Perez KR, Miranda A, Riske KA. Comparative Study Of The Mechanism Of Action Of The Antimicrobial Peptide Gomesin And Its Linear Analogue: The Role Of The Beta-Hairpin Structure. Biochimica et Biophysica Acta. 2015;1848(10 Pt A):2414-21.

56. Domingues TM, Riske KA, Miranda A. Revealing The Lytic Mechanism Of The Antimicrobial Peptide Gomesin By Observing Giant Unilamellar Vesicles. Langmuir. 2010;26(13):11077-84.
57. Zwaal RFA, Comfurius P, Bevers EM. Surface Exposure Of Phosphatidylserine In Pathological Cells. Cellular and Molecular Life Sciences. 2005;62(9):971-88.
58. Dowhan W. Molecular Basis For Membrane Phospholipid Diversity: Why Are There So Many Lipids? Annual Review of Biochemistry. 1997;66(1):199-232.
59. Paredes-Gamero EJ, Casaes-Rodrigues RL, Moura GE, Domingues TM, Buri MV, Ferreira VH, et al. Cell-Permeable Gomesin Peptide Promotes Cell Death By Intracellular Ca(2+) Overload. Molecular Pharmaceutics. 2012;9(9):2686-97.
60. Fazio MA, Jouvensal L, Vovelle F, Bulet P, Miranda MT, Daffre S, et al. Biological And Structural Characterization Of New Linear Gomesin Analogues With Improved Therapeutic Indices. Biopolymers. 2007;88(3):386-400.
61. Fazio MA, Oliveira VX, Jr., Bulet P, Miranda MT, Daffre S, Miranda A. Structure-Activity Relationship Studies Of Gomesin: Importance Of The Disulfide Bridges For Conformation, Bioactivities, And Serum Stability. Biopolymers. 2006;84(2):205-18.
62. Deplazes E, Chin YK-Y, King GF, Mancera RL. The Unusual Conformation Of Cross-Strand Disulfide Bonds Is Critical To The Stability Of B-Hairpin Peptides. Proteins: Structure, Function, and Bioinformatics. 2020;88(3):485-502.
63. Koziara KB, Stroet M, Malde AK, Mark AE. Testing And Validation Of The Automated Topology Builder (Atb) Version 2.0: Prediction Of Hydration Free Enthalpies. Journal of Computer-Aided Molecular Design. 2014;28(3):221-33.
64. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. Gromacs 4: Algorithms For Highly Efficient, Load-Balanced, And Scalable Molecular Simulation. Journal of Chemical Theory and Computation. 2008;4(3):435-47.
65. Schmid N, Eichenberger A, Choutko A, Riniker S, Winger M, Mark A, et al. Definition And Testing Of The Gromos Force-Field Versions 54a7 And 54b7. European Biophysics Journal. 2011;40(7):843-56.
66. Poger D, Van Gunsteren WF, Mark AE. A New Force Field For Simulating Phosphatidylcholine Bilayers. Journal of Computational Chemistry. 2010;31(6):1117-25.
67. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction Models For Water In Relation To Protein Hydration. In: Pullman B, editor. Intermolecular Forces. The Jerusalem Symposia on Quantum Chemistry and Biochemistry. 14: Springer Netherlands; 1981. p. 331-42.
68. Tironi IG, Sperb R, Smith PE, Gunsteren WFv. A Generalized Reaction Field Method For Molecular Dynamics Simulations. Journal of Chemical Physics. 1995;102(13):5451-9.
69. Berendsen HJC, Postma JPM, van Gunsteren WF, Dinola A, Haak JR. Molecular Dynamics With Coupling To An External Bath. Journal of Chemical Physics. 1984;81(8):3684-90.
70. Nosé S. A Unified Formulation Of The Constant Temperature Molecular Dynamics Methods. Journal of Chemical Physics. 1984;81(1):511-9.
71. Parrinello M, Rahman A. Polymorphic Transitions In Single Crystals: A New Molecular Dynamics Method. Journal of Applied Physics. 1981;52(12):7182-90.
72. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. Lincs: A Linear Constraint Solver For Molecular Simulations. Journal of Computational Chemistry. 1998;18(12):1463-72.
73. Miyamoto S, Kollman PA. Settle: An Analytical Version Of The Shake And Rattle Algorithm For Rigid Water Models. Journal of Computational Chemistry. 1992;13(8):952-62.
74. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. The Weighted Histogram Analysis Method For Free-Energy Calculations On Biomolecules. I. The Method. Journal of Computational Chemistry. 1992;13(8):1011-21.

75. Hub JS, de Groot BL, van der Spoel D. g_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. Journal of Chemical Theory and Computation. 2010;6(12):3713-20.
76. Michaud-Agrawal N, Denning EJ, Woolf TB, Beckstein O. Mdanalysis: A Toolkit For The Analysis Of Molecular Dynamics Simulations. Journal of Computational Chemistry. 2011;32(10):2319-27.
77. Hunter JD. Matplotlib: A 2d Graphics Environment. Computing in Science & Engineering. 2007;9(3):90-5.
78. Troeira Henriques S, Lawrence N, Chaousis S, Ravipati AS, Cheneval O, Benfield AH, et al. Redesigned Spider Peptide With Improved Antimicrobial And Anticancer Properties. ACS Chemical Biology. 2017;12(9):2324-34.
79. Cheng JTJ, Hale JD, Elliott M, Hancock REW, Straus SK. The Importance Of Bacterial Membrane Composition In The Structure And Function Of Aurein 2.2 And Selected Variants. Biochimica et Biophysica Acta - Biomembranes. 2011;1808(3):622-33.
80. Bernardi RC, Melo MCR, Schulten K. Enhanced Sampling Techniques In Molecular Dynamics Simulations Of Biological Systems. Biochimica et Biophysica Acta - General Subjects. 2015;1850(5):872-7.
81. Abrams CB, Giovanni. Enhanced Sampling In Molecular Dynamics Using Metadynamics, Replica-Exchange, And Temperature-Acceleration. Entropy. 2014;16(1):163-99.
82. Marzuoli I, Margreitter C, Fraternali F. Lipid Head Group Parameterization For Gromos 54a8: A Consistent Approach With Protein Force Field Description. Journal of Chemical Theory and Computation. 2019;15(10):5175-93.
83. Hwang PM, Zhou N, Shan X, Arrowsmith CH, Vogel HJ. Three-Dimensional Solution Structure Of Lactoferricin B, An Antimicrobial Peptide Derived From Bovine Lactoferrin. Biochemistry. 1998;37(12):4288-98.
84. Tolkien JRR. The Hobbit. New York: Ballantine Books; 1937.

6.0 Supporting Information

Table S1: Replica exchange regimes tested in Sections 3.2 and 3.3 detailing the number of replicas in each regime and the interaction scaling factor (SF) used in each replica.

		Peptide-environment scaling factor									
Regime Name	Number of replicas	0 (Ground)	1	2	3	4	5	6	7	8	9
Max SF 0.8	3	1.0	0.9	**0.8**							
Max SF 0.6	5	1.0	0.9	0.8	0.7	**0.6**					
Max SF 0.4	7	1.0	0.9	0.8	0.7	0.6	0.5	**0.4**			
Max SF 0.2	10	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.25	**0.2**

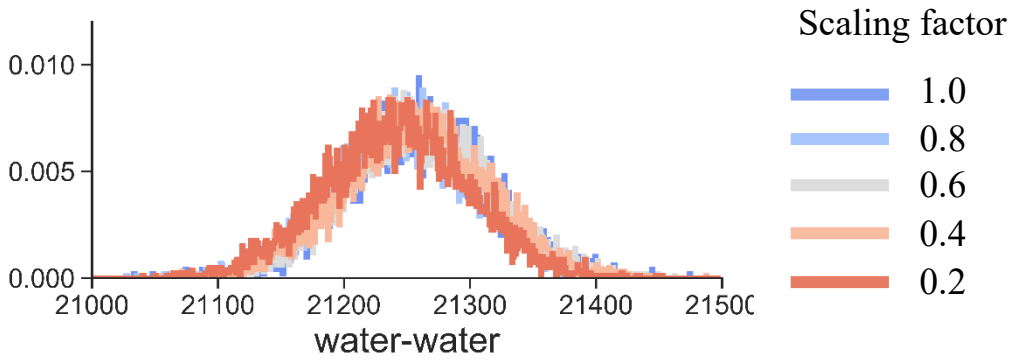


Figure S1: Normalised histogram of the number of water-water hydrogen bonds (HBs) in unscaled (SF = 1.0) and scaled simulations with peptide-environment interaction SFs of 0.8, 0.6, 0.4, and 0.2. (Table 1: cUS-2.6 and SS-0.8 → 0.2).

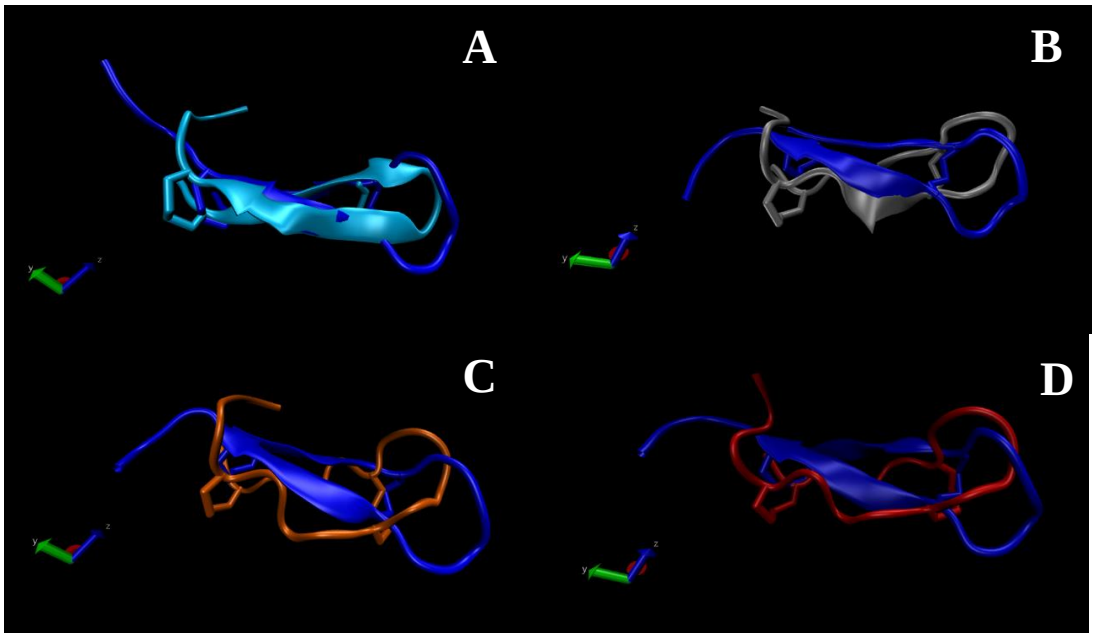


Figure S2: Structure of Gm in a simulation with scaled peptide-environment interactions compared to its structure in an unscaled simulation (blue) (Table 1: cUS-2.6 and SS-0.8 → SS-0.2). The peptide-environment interaction SFs were A: 0.8, B: 0.6, C: 0.4, and D: 0.2. In all simulations the peptide was held at a COM distance of 2.6 nm to the membrane. The membrane is not shown in these images for clarity. The structures shown here are not necessarily representative of the conformation of Gm across the simulation trajectory, and are only examples of the deformations that were observed to occur.

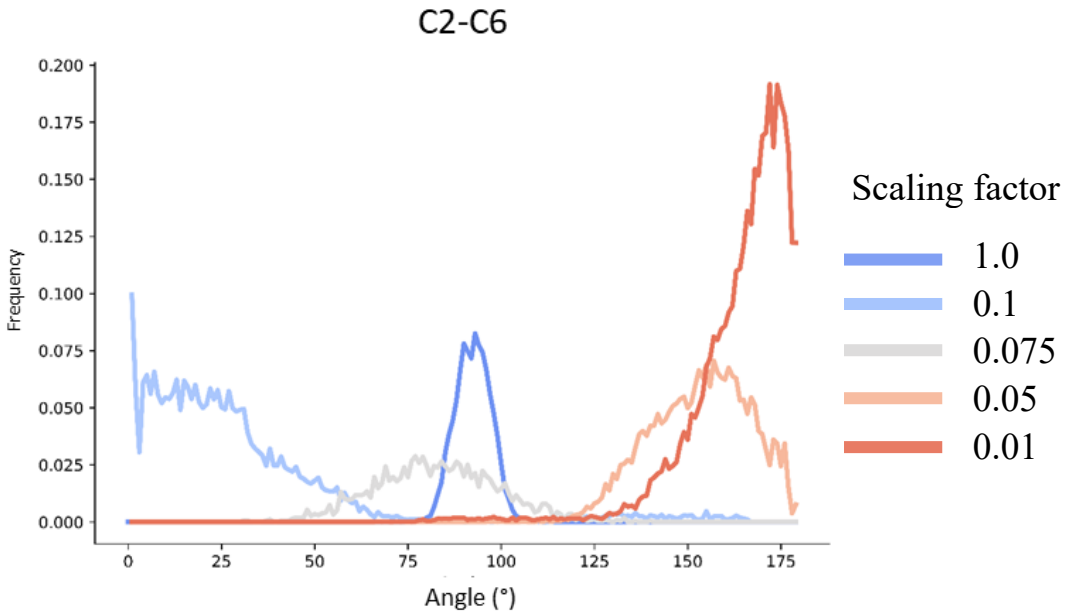


Figure S3: Frequency- and angle-normalised distributions of the C2-C6 vector of Gm, held at a COM distance of 2.2 nm (Table 1: SS2.2-0.2 → SS2.2-0.01). The peptide-environment interaction SFs tested were 0.1, 0.075, 0.05, and 0.01.

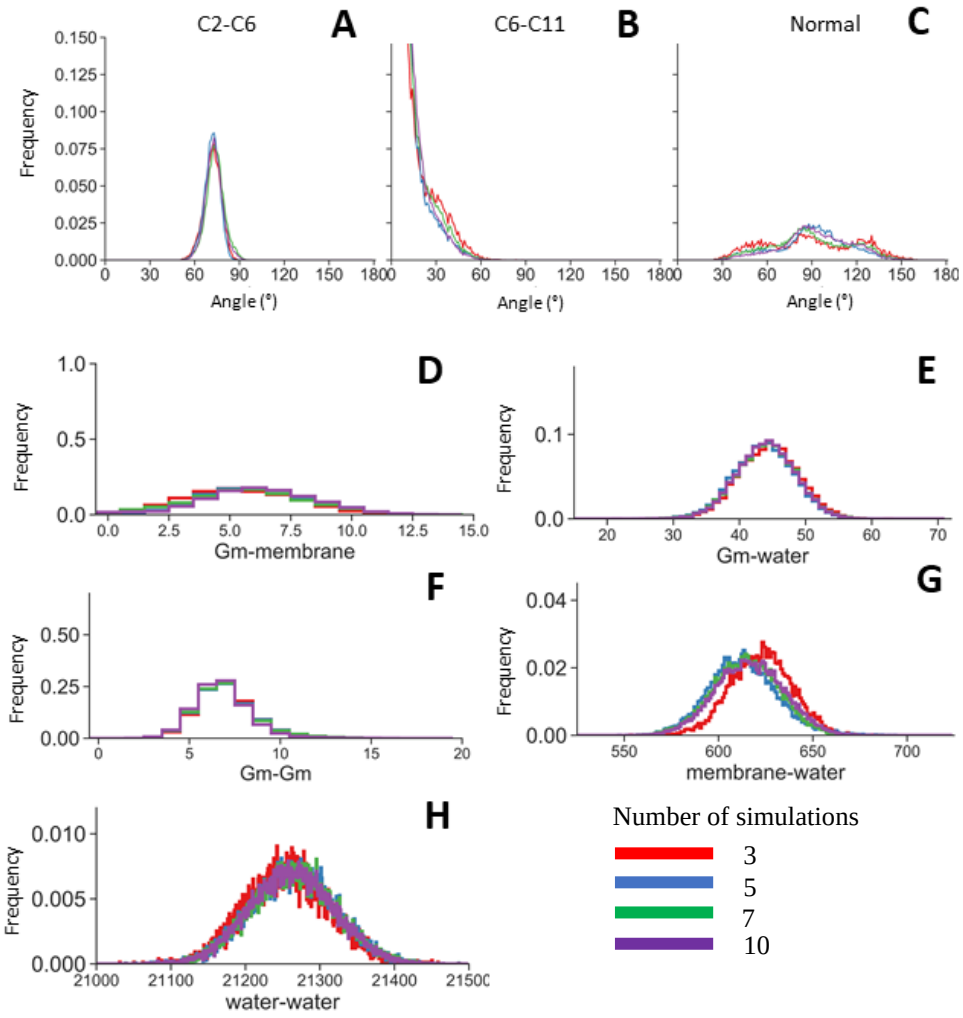


Figure S4: Influence of additional simulation data on angle and HB distributions. A-C: Frequency- and angle-normalised distributions of the C2C6, C6C11, and normal vectors. Each line represents the combined data from the specified number of independent 50 ns simulations of conventional MD (Table 1: Cr-1 → Cr-10). D-H: Normalised histograms of the number of HBs between the simulation system components of Gm-membrane, Gm-water, Gm-Gm, membrane-water, and water-water.

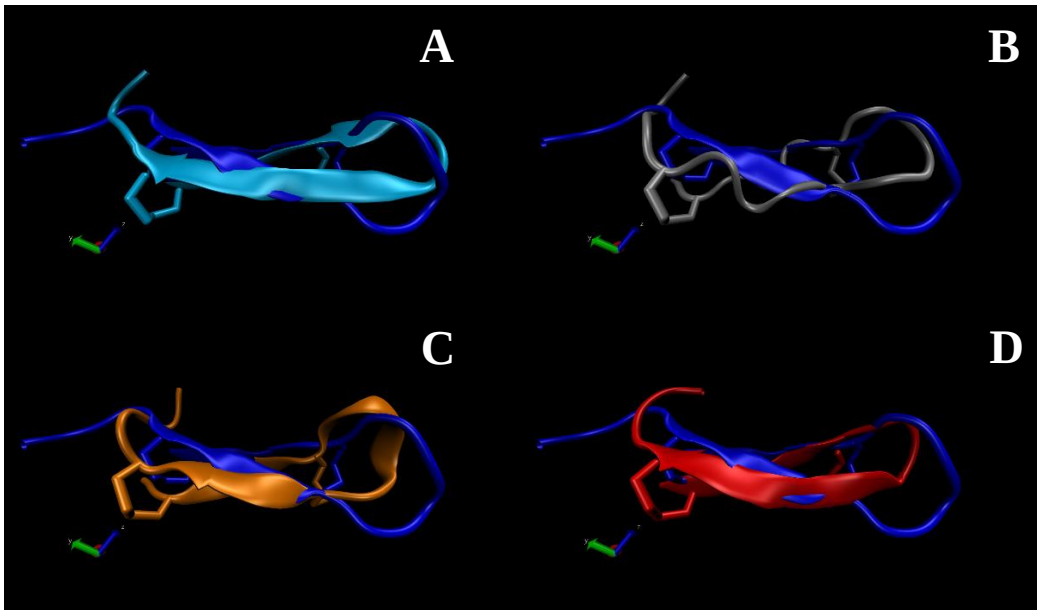


Figure S5: Comparison of the structure of Gm in unscaled (blue) and ground replicas of the replica exchange regime simulations with Max SFs A: 0.8, B: 0.6, C: 0.4, D: 0.2 (Table 1: R-0.8 → R-0.2). In all simulations the peptide was held at a COM distance of 2.6 nm to the membrane. The membrane is not shown in these images for clarity. The structures shown here are not necessarily representative of the conformations of Gm sampled across the simulation trajectory, and are only examples of the deformations that can occur.

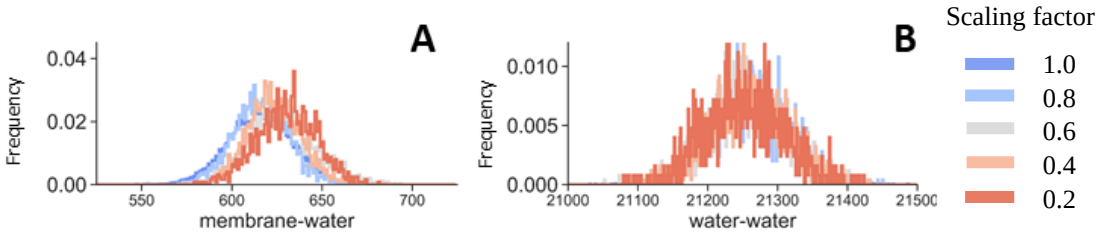


Figure S6: Normalised histograms of the number of A: membrane-water HBs, and B: water-water HBs, calculated in the ground replica of the replica exchange regime simulations with Max SFs of 0.8, 0.6, 0.4, and 0.2. (Table 1: R-0.8 → R-0.2).

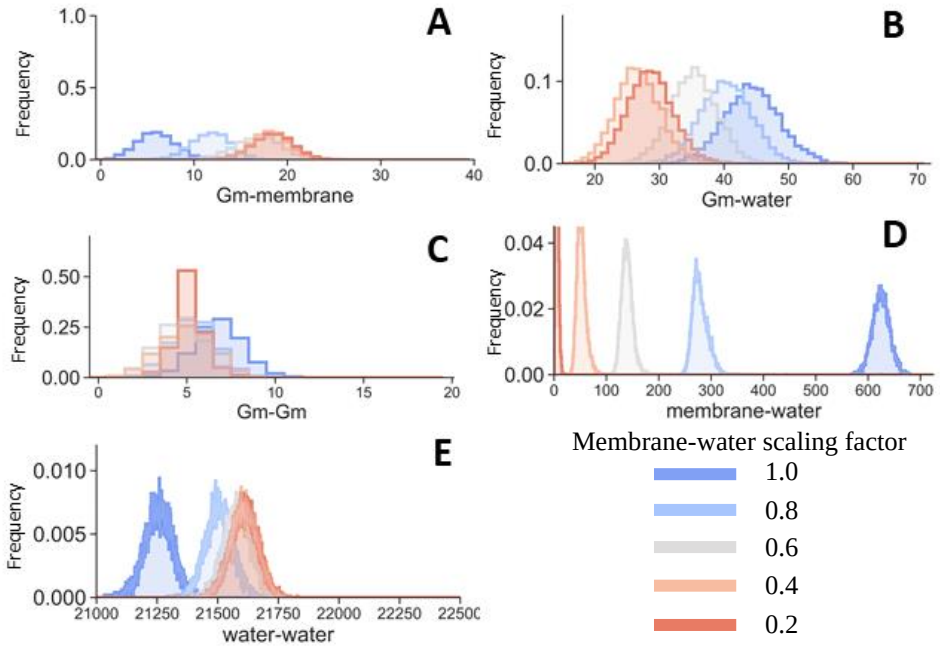


Figure S7: Normalised histograms of the number of A: Gm-membrane, B: Gm-water, C: Gm-Gm, D: membrane-water, and E: water-water HBs, from scaled simulations with membrane-water interaction SFs of 0.8, 0.6, 0.4, and 0.2 (Table 1: sMW-0.8 → sMW-0.2). Peptide-environment interactions are unscaled

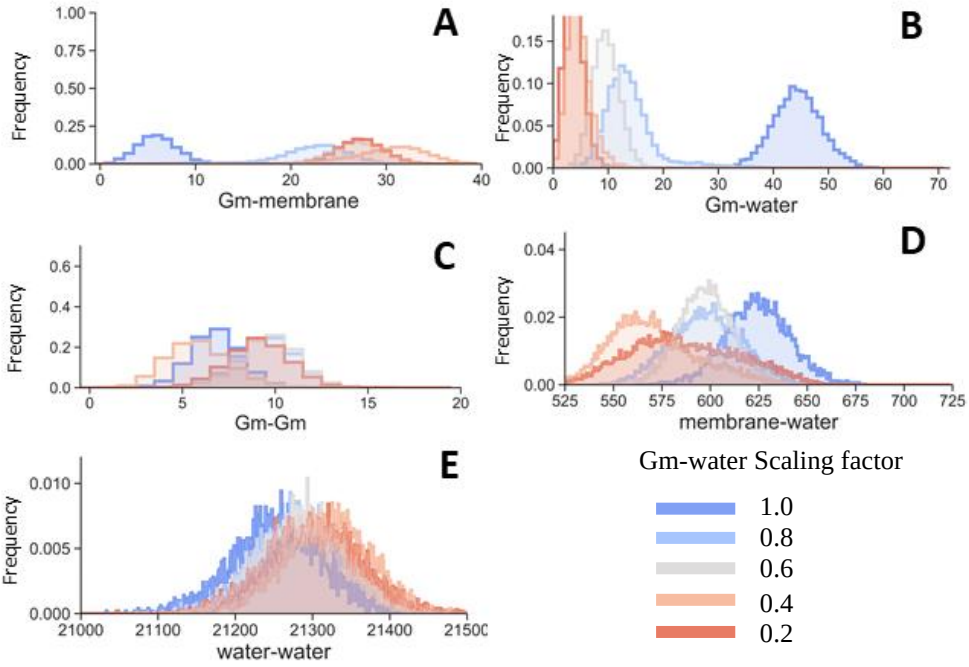


Figure S3: Normalised histograms of the number of A: Gm-membrane, B: Gm-water, C: Gm-Gm, D: membrane-water, and E: water-water HBs, from scaled simulations with unscaled

peptide-membrane interactions (SF = 1.0) and with peptide-water interactions scaled by 0.8, 0.6, 0.4, and 0.2. (Table 1: sUE-0.8 → sUE-0.2)