# Conformations of myoglobin-derived peptides at the air-water interface

David L. Cheung*

*School of Chemistry, National University of Ireland Galway, Galway, Ireland*

E-mail: david.cheung@nuigalway.ie

**Abstract**

The conformational change exhibited by proteins at liquid interfaces, such as the air-water and oil-water interfaces, has long been of interest, both for understanding protein structure outside of native environments and for applications in areas including food technology and pharmaceuticals. Using molecular simulation this paper studies the conformations of two peptides derived from myoglobin, for which the emulsification behaviour has been studied. Both peptides were found to readily adsorb onto the air-water interface, with one of these (experimentally the more effective stabiliser) adopting a flat, extended conformation, with the other peptide remaining close its solution conformation.

## Introduction

Containing both hydrophobic and hydrophilic amino acids proteins are intrinsically amphiphilic molecules and readily adsorb onto hydrophobic interfaces, such as the air-water or oil-water interfaces. This is typically associated with conformational change,[1] as the protein

*To whom correspondence should be addressed

rearranges to allow the core hydrophobic residues to come into contact with the hydrophobic medium, and so is typically avoided in most biological situations. For a small number of proteins, however, adsorption at interfaces is a key part of their function and these have evolved a number of strategies for interfacial adsorption. These include the amphiphilic surface structure of hydrophobins[2] or specific conformational changes seen in proteins such as ranaspumin,[3] latherin,[4] or lipase.[5] Non-surfactant proteins are also found to adsorb onto liquid interfaces and these are commonly used as foam or emulsion stabilisers,[6] with milk and whey proteins being used in food and other consumer products. Alongside naturally occurring proteins, the behaviour of a number of designed peptides, such as the LK-peptides[7] and AM1,[8] at interfaces have been investigated. As well as being important in identifying and designing proteins for applications, study of protein conformational change at liquid interfaces also addresses fundamental questions regarding protein function outside of native environments.

Due to this scientific and technological interest the behaviour of proteins at liquid interfaces has been the subject of extensive experimental investigation. Tensiometry and rheological measurements have been extensively applied to the study of proteins at liquid interfaces; while these cannot give direct insight into protein structure changes in interfacial properties they can be used to indirect infer changes in protein conformation, e.g. changes in viscosity and elasticity through the formation of interfacial gels or crystals. The tertiary structure of proteins, for example the thickness of adsorbed layers, has been examined through X-ray[9] and neutron reflectivity.[10] In some cases this layer thickness can be related to changes in protein conformation; for example the change in thickness of the protein Rsn-2[3] at the air-water interface measured through neutron reflectivity is commensurate with a change in structure from a closed to open conformation at the air-water interface. Secondary structure can be investigated using circular dichroism and IR spectroscopy. In particular the development of refractive indexed matched emulsion circular dichoism (RIME CD)[11] and synchrotron radiation circular dichroism (SRCD)[12] has allowed for the quantitative determination of protein

secondary structure at emulsion interfaces. Application of this to a number of different proteins[13,14] has shown that secondary structure change of proteins at oil-water interfaces depends on the nature of the oil-phase and on the specific protein sequence.

Alongside experimental measurements molecular simulation has been used to investigate the adsorption of proteins on liquid interfaces. Due to the system sizes involved, these have typically involved the investigation of systems containing single proteins, probing the initial stages of interfacial adsorption or in determining adsorption free energies. Both atomistic[15–18] and coarse-grain models[19,20] have been used, the latter being more computationally efficient and due to their simpler parameterization capable of investigating the link between protein surface structure and interfacial behaviour.[20] Atomistic models by contrast are capable of modelling protein structural change at interfaces, giving some insight into protein function such as biomineralization[21] and catalysis[22] at interfaces. The combination of atomistic simulations and neutron reflectivity has proven to be increasingly common,[23,24] as the simulations can give direct insight into protein structure that can be used to rationalise the NR measurements.

The limited timescales that simulation is able to address has limited their utility in modelling large scale protein conformational change at interfaces. As changes in conformation related to function, such as the lid opening in lipases,[25] occur over times greater than tens of nanoseconds simulations in excess of 100 ns are typically required to study these. Changes in protein conformation can also be rare events, for which a significant waiting time is necessary before they occur. Novel simulation methods, such as replica exchange[26] or accelerated molecular dynamics,[27] have often been used in studies of proteins in bulk solution, and have recently been applied to the study of EAS[28] and Rsn-2 at interfaces. Similar methods have also been applied to the investigation of proteins and peptides at soft surfaces, such as self-assembled monolayers.[29,30]

As its structure has long been known myoglobin[31] has been regarded as a model protein. It has a well defined tertiary structure consisting of eight alpha helices (Figure 1(a)), enclosing

the functional heme group. The primary function of myoglobin is the transport of molecular oxygen through its binding to the heme group. This is located near the protein surface, allowing for contact between ligand and heme group, although a conformational change occurs upon oxygen binding. This structural change is aided by the flexible structure of myoglobin due to its lack of disulphide bonds. Many of its helices are amphipathic leading to adsorption at hydrophobic interfaces and its interfacial and emulsifying properties have been the subject of a number of studies.[14,32,33] Due to its lack of disulphide bonds it is less stable at interfaces then other commonly studied proteins, such as lysozyme, which was confirmed using synchrotron radiation circular dichroism measurements.[14] In order to obtain a better understanding of the relation between the protein structure and interfacial function Poon *et al* studied apo-myoglobin (myoglobin without the heme group) and a number of peptides derived from this.[32,34] Using cyanogen bromide to cleave after methionine residues, two easily purified peptides (1-55 and 56-131) were found. Within the full protein the second peptide largely encompasses the heme group, which is bonded to the N$\epsilon$ atom in the imidazole sidechain of residue 93. This helps stabilise the structure of this section of the protein, in which the helices are less amphipathic than in the first peptide (1-55). Comparison between these and apo-myoglobin, showed that peptide1-55 was an effective emulsifier whereas peptide56-131 performed more poorly.[32]

As these are fragments of a larger protein, which in of itself has not evolved to function at liquid interfaces, these are interesting test cases for the studying the relationship between protein structure and interfacial conformation. They are more complex than *de novo* designed sequences, such as the LK peptides, and so may be expected to show more complex behaviour, more applicable to naturally occurring peptides and proteins. Comparison between these two peptides shows that they have similar proportions of hydrophobic amino acids (45% for peptide1-55, 43% for peptide56-131) but the distribution of hydrophobic amino acids within these differs with peptide1-55 exhibiting a more amphiphilic character than peptide56-131 (Figure 1(b)), which was suggested as the cause of their differing

emulsification properties.[32]

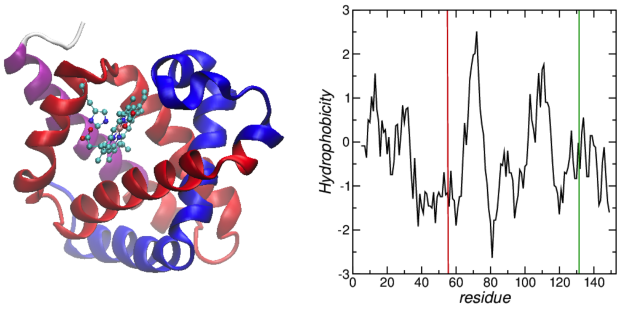



Figure 1: (a) Structure of myoglobin (1MBN). Peptide1-55 and 56-131 shown in blue and red respectively, heme group and histidine sidechain represented using CPK. (b) Kyte-Doolittle[35] hydropathy plot for myoglobin, vertical lines denote peptide1-55 and 56-131. Data generated using the ExPASy webserver[36] (http://web.expasy.org/cgi-bin/protscale/protscale.pl)

In this paper atomistic molecular dynamics simulations have been used to investigate the structures of these two peptides at the air-water interface. While they both readily adsorb onto the interface, peptide1-55 can undergo a significant tertiary structure change upon interfacial adsorption. By contrast peptide56-131, which is the poorer emulsifier, only exhibits only a small change in structure. Replica exchange with solute tempering (REST) simulations reveal that both these peptides can adopt a number of distinct conformations at the interface; for peptide1-55 adopts a flat conformation, maximising its interfacial area and potentially leading to favourable interfaces with other adsorbed molecules. The stability of these different conformations relies on the interplay between interfacial free energy and intramolecular interactions.

# Simulation details

## System and methodology

The starting configurations for the peptides in bulk water were taken from the myoglobin crystal structure (pdb: 1MBN[31]). Initially these fragments were placed in a cubic water box. The charges on the ionisable residues were set appropriately for pH 7. $Na^+$ or $Cl^-$ ions added

to neutralise the system (no additional counter-ions were added). Peptide conformations from the bulk simulations were used as starting structures for the air-water simulations; for both peptide three separate air-water simulations were performed, using conformations spaced 10 ns apart. These were placed in the centre of a cuboidal water box (size 7 nm×7 nm×5 nm) and then the box was extended up to 15 nm in the z direction.

In order to improved the sampling of peptide conformations at the air-water interface replica exchange with solute tempering[37,38] (REST) is used. This is a modification of replica exchange[26] where only a subset of the system, in this case the peptide, is run at different temperatures, which corresponds to scaling the peptide-peptide and peptide-solvent interactions by factors of $T_i/T$ and $\sqrt{T_i/T}$, where. $T_i$ is the temperature of each replica. Compared to standard REMD this is needs fewer replicas to cover a given temperature range. For each peptide one REST simulation was performed (taken from the end point of the first NVT simulation). In total six simulations were performed for each peptide (one in bulk solution, three at the air-water interface at 298 K, one at the air-water interface at 350 K, and one REST simulation). Across these different simulations the total simulation time was 6.7 $\mu s$. It should be noted that in recent years a number of other methods for accelerating the sampling of protein conformations, such as metadynamics or accelerated dynamics, have been developed.[39] While these can be competitive if not more efficient than replica exchange, they typically require knowledge of the change in protein structure (e.g. through the collective variables used to bias the dynamics). In this study the different interfacial conformations were not known *a priori* and so REST was a more suitable choice.

The simulations were performed using the Gromacs molecular dynamics package.[40,41] For the REST simulations version 4.6.3 using a modified version of PLUMED[38,42] was employed, with Gromacs 5.0 used for the other simulations. Standard gromacs tools were used to create the simulation input files. The Gromos54a7 force field[43] was used along with SPC water.[44] Van der Waals and real-space electrostatic interactions were truncated at 10 Å. Long-range electrostatic interactions were accounted for using PME;[45] reciprocal space grids

of 40×40×40 cells and 44×44×96 cells were used for the bulk and air-water simulations respectively.

Bulk simulations were performed in the NpT-ensemble and the the air-water simulations in the NVT-ensemble. In both cases temperature was controlled using a velocity rescaling algorithm[46] (relaxation time 0.1 ps) and for the bulk simulations pressure was controlled using the Parrinello-Rahman barostat, with relaxation time 2 ps. Bond lengths were constrained using the LINCS algorithm.[47] All systems were energy minimized using the steepest descents algorithm, followed by a short NVT simulations (20 ps) with the heavy atoms in the peptide constrained to their initial positions. Following this short NVT, and for the bulk systems NpT simulations (both 20 ps), without constraints on heavy atoms were performed. For the REST simulations 6 replicas were used, with scaling factors $\lambda$ =1.0 (298 K), 0.968 (307.74 K), 0.938 (317.80), 0.908 (328.19 K), 0.879 (338.92 K), and 0.851 (350 K). This range of temperatures was chosen to encompass the major changes in peptide conformation at the air-water interface, which for both peptides occurs below 350 K (see below). In acceptance rates were typically 20%-30% (Table S1). In order to monitor the convergence of the REST simulations the running average of the populations of the different states was calculated (Fig. S1). For both peptides the populations of the most occupied states had reached constant values prior to the data gathering stage of the simulation (last 100 ns).

## Analysis

The simulations were analysed using a combination of standard gromacs tools, VMD, and custom written python scripts, using the MDAnalysis library.[48] Potential energy contributions (Table 2) were found using the rerun option of mdrun. Representative structures for the different interfacial conformations were determined using the g_cluster tool using the Gromos criteria,[49] with $R_{cut} = 3$ Å. The position of the air-water interface, for determining the protein and residue centre-of-mass separations, was determined from the Gibbs dividing surface, found following Vink *et al.*[50] To estimate the surface area a grid of points is placed

on the air-water interface, defined by the Gibbs surface with the area being estimated from the number of grids points that were within a cut-off distance from a protein atom, given by the van der Waals radius+1.4 Å(the van der Waals radius of water oxygen). The partition free energy was found by summing the water-air transfer free energies[51] for each amino acid whose side-chain centre of mass was in the air ($z > z_{interface}$). Simulation snapshots were generated using VMD,[52] with the STRIDE algorithm[53] used for secondary structure assignment. Python scripts used in the analysis may be obtained on request from the author or via the website (http://www.nuigalway.ie/softinterfaces).

# Results

## Solution structures of peptides 1-55 and 56-131

For peptide1-55 the initial extended structure, consisting primarily of two $\alpha$-helices, rapidly contracts into a globular structure. The principal moments of the gyration tensor ($R_G^i$) are 7.23 Å, 5.95 Å, and 5.50 Å, showing that it adopts an approximately cylindrical shape. From the simulation snapshots and secondary structure analysis the two main $\alpha$-helices (residues 4-19 and 22-35) remain largely intact, with the second helix shrinking slightly, but the angle between the decreases so they lie close to parallel. The third helical region in the initial structure (residues 36-42) disappears with a new region of $\alpha$-helix appearing (residues 44-48). Due to this rearrangement of helical residues and the change in tertiary structure the number of residues in $\alpha$-helices decreases from 36 in the initial structure to 31, with a slight increase in random coil to accommodate the rearrangement of the helices. Overall the main secondary structure components of peptide1-55 in bulk water are $\alpha$-helix 56.1 %, turn 17.4 %, and random coil 24.9 %.

Due to its larger size structural change for peptide56-131 occurs over a longer timescale. While for the first 90 ns of the simulation the radius of gyration and its two largest eigenvalues remain largely constant, these show significant fluctuations. After 90 ns the radius of gyration

and its largest eigenvalue suddenly drop, indicating a transition to a shorter, more compact structure. This the radius of gyration then remains constant with smaller fluctuations. Two of the helical regions (residues 59-95 and 101-119) remain largely intact throughout the simulation. The central helix (residues 81-95), however, decreases in size with the residues towards the C-terminus end converting to turn. This central helical region is significantly less hydrophobic than the outer two, which may reduce the stability of the helix in solution. In the full protein the link to the heme group is also contained within this region and removal of this may also reduce the stability of the helix. Overall the main secondary structure components of peptide56-131 in bulk water are $\alpha$-helix 63.6 %, turn 14.5 %, and random coil 21.3 %. This represents a significant decrease in the $\alpha$-helix content compared to the starting conformation (taken directly from the full protein), for which the $\alpha$-helix content is $\sim 80$ %.

## Both peptides 1-55 and 56-131 readily adsorb onto air-water interface

When placed in the centre of the water slab peptide1-55 rapidly adsorbs onto the air-water interface (Fig. 3(a)). For all three replicate runs this occurs within approximately 10 ns, which is comparable to the time expected for the peptide to reach the interface under diffusive motion. The diffusion coefficient found from the bulk simulations is $D = 8.22 \times 10^{-3}$ nm$^2$ ps$^{-1}$ so the time expected for the peptide to reach the interface (given a slab width $L_z = 5$ nm) under diffusive motion is

$$t_{diff} = \frac{(L_z/2)^2}{D} = \frac{(2.5 \text{ nm})^2}{8.22 \times 10^{-3} \text{ nm}^2 \text{ ps}^{-1}} = 7.6 \text{ ns}.$$

This suggests that there is little or no energy barrier to interfacial adsorption. Due to the small number (three) of simulations performed this should not be regarded as a prediction of the adsorption rate for this peptide, rather simply to observe that it falls into the timescales

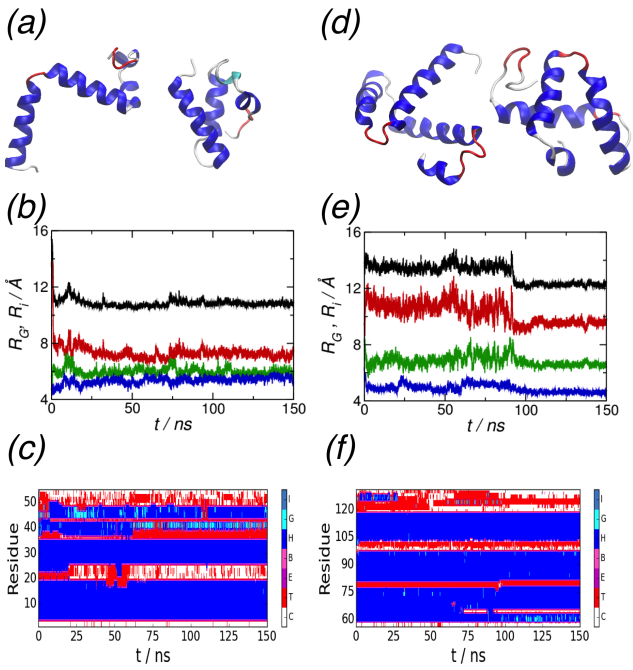



Figure 2: Structures formed by peptide1-55 and 56-131 in bulk solution (a,d) Initial (left) and final (right) solution structures for peptide1-55 (a) and 56-131 (d). (b) and (e) Radius of gyration and principal moments of the gyration tensor. $R_G$, $R_G^1$, $R_G^2$, and $R_G^3$ denoted by black, red, green, and blue lines respectively. (c) and (f) Secondary structure distributions. blue denotes $\alpha$-helix, red turn, cyan $3_{10}$-helix, and white random coil.

that would be expected for purely diffusive motion. Additionally no attempt is made to determine the adsorption free energy from these simulations. Free energy calculations, using methods such as steered molecular dynamics,[54] adaptive biasing force,[55] or metadynamics,[56] may be used to calculate this; however, this is outside the scope of the present study. After this the behaviour of the peptide for the first ∼200 ns of the simulation is similar for all the replicate runs with the peptide sitting approximately 7 Åfrom the interface. Analysis of the individual residue positions (Fig. 3(b)) shows that that N-terminus of the peptide (including the first $\alpha$-helix) and the loop joining the two remaining helices lie closest to the interface, with the remaining helices remaining more distant. Throughout this period both the tertiary and secondary structures of the peptide remain largely unchanged, similar to the solution structure.

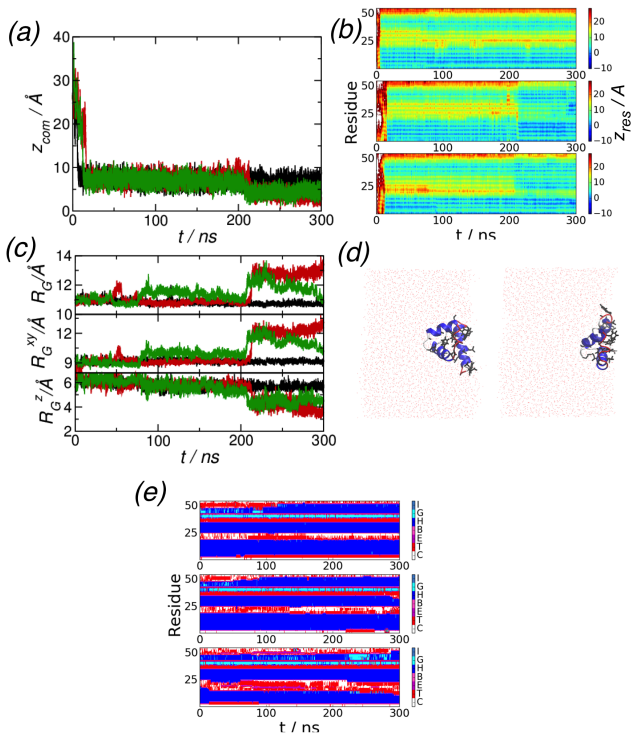



Figure 3: Position and structure of peptide1-55 at an air-water interface at 298 K: (a) Separation between peptide centre of mass and air-water interface ($z_{com}$). First, second, and third replicate runs denoted by black, red, and green lines respectively. (b) Separation between residue centre-of-mass and air-water interface ($z_{res}$). (c) Radius of gyration (top), $R_G^{xy}$ (middle), and $R_G^z$ (bottom). Symbols as in (a). (d) End simulation snapshots for runs 1 (left) and 2 (right). Hydrophobic sidechains shown in black. (e) Secondary structure distributions for run 1 (top), run 2 (middle), and run 3 (bottom). Blue denotes $\alpha$-helix, red turn, cyan $3_{10}$-helix, and white random coil.

At longer times differences between the different replicates are seen with the peptide centre of mass moving closer to the interface ($\sim$ 5 Å) after about 200 ns in two of the simulations. This shift in centre of mass position occurs due to a change in the peptide structure with the two $\alpha$-helices moving onto the interface. This results in the peptide adopting a flat conformation. This change in arrangement may be seen in the radius of gyration ($R_G$) and its parallel ($R_G^{xy}$) and perpendicular components ($R_G^z$), which can be identified with the peptide size in the plane of the interface and in the $z$-direction respectively. By contrast the first run, which showed no change in the $z_{com}$ throughout the simulation, $R_G$ and its parallel and perpendicular components remain approximately constant, with values

close to those of the solution structure. The differences between the simulation runs suggests that peptide1-55 can adopt multiple conformations at the air-water interface. This change in conformation can also be seen in simulation snapshots (Fig. 3(d)). In the flat conformation the rearrangement of the peptide allows for more of the hydrophobic sidechains to move out of the water, in contrast to the compact state where these are largely confined to the peptide core.

As in bulk solution the secondary structure is dominated by three $\alpha$-helical segments, approximately residues 6-18, 26-35, and 44-52, with small segments of turn. In all three replicate runs (Fig. 3(e)) there is a short segment, approximately residues 44 to 48, switches between turn and $3_{10}$-helix (typically 55 % of the time turn, 40 % $3_{10}$-helix). This section corresponds to a section of $\alpha$-helix found in the crystal structure of myoglobin missing in the solution structure of peptide 1-55. While there is some variation between the different runs the overall amount of ordered structure increases compared to the solution structure, with proportion of random coil decreasing from $\sim$25 % to $\sim$ 20 % (Table 1).

For peptide56-131 more variation is seen in the time needed for adsorption to the air-water interface. In two of the replicate runs this is again very rapid ($\sim$ 10 ns) which is again comparable to the diffusive timescale ($D = 5.07 \times 10^{-3}$ nm$^2$ ps$^{-1}$ giving $t_{diff} = 12.6$ ns), while for the remaining run adsorption takes significantly longer ($\sim$ 50 ns). Again this should not be regarded as quantitative prediction of the adsorption rate, rather a suggestion that adsorption is most likely purely diffusion limited. Once adsorbed the location of the peptide relative to the interface also varies between the replicate runs; either it lies close to the interface (with $z_{com} \sim 5$ Å), it lies far from the interface ($\sim$ 15 Å) or it initially lies far from the interface before switching to a closer location. In all cases the peptide adsorbs through its centre (approximately residues 75-105), with the adsorption of the N-terminus end of the peptide to the interface begin associated with the closer adsorption seen in runs 1 and 3. For all the replicate runs the C-terminus end of the peptide remains further away from the interface, consistent with the less hydrophobic character of this end of the peptide

(Fig. 1(b)).

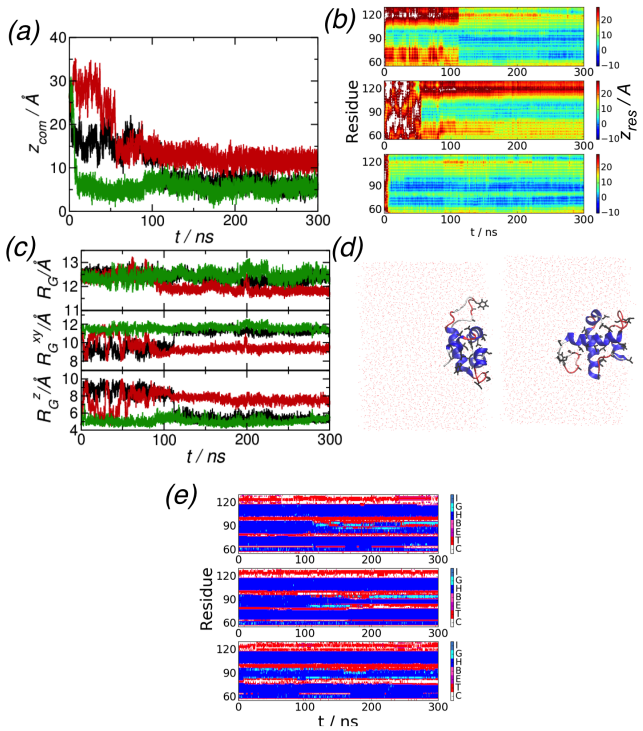



Figure 4: Position and structure of peptide56-131 at an air-water interface at 298 K: (a) Separation between peptide centre of mass and air-water interface ($z_{com}$). First, second, and third replicate runs denoted by black, red, and green lines respectively. (b) Separation between residue centre-of-mass and air-water interface ($z_{res}$). (c) Radius of gyration (top), $R_G^{xy}$ (middle), and $R_G^z$ (bottom). Symbols as in (a). (d) End simulation snapshots for runs 1 (left) and 2 (right). Hydrophobic sidechains shown in black. (e) Secondary structure distributions for run 1 (top), run 2 (middle), and run 3 (bottom). Blue denotes $\alpha$-helix, red turn, cyan $3_{10}$-helix, and white random coil.

Unlike peptide1-55 adsorption of peptide56-131 close to the interface is not associated with significant changes to the radius of gyration (Fig. 4(c)), with $R_G$ and its parallel and perpendicular components remaining close to their solution values across the simulation. More variation in $R_G$ is seen in the second run in which the peptide lies further from the interface. From the relative invariance of $R_G$ it can be deduced that the changes in $R_G^{xy}$ and $R_G^z$ correspond to changes in the orientation of the peptide, rather than its structure. This can also be seen in the eigenvalues of the gyration tensor which remain close to the solution values throughout the simulation (Fig. S2). In both the first and second runs the peptide

initially attaches normal to the interface, with the protein reorienting itself after $\sim$ 120 ns for the first run. This difference in orientation can be seen in the simulation snapshots (Fig. 4(d)). The secondary structure changes more upon interfacial adsorption than peptide1-55, with the helix in contact with the interface (residues 82-96) becoming shorter, which leads to an overall decrease in the proportion of $\alpha$-helix (Table 1). Molecular dynamics simulations of isolated myoglobin helices shows that this helix is relatively unstable.[57]

Table 1: Secondary structure components for peptides in bulk solution and at air-water interface at 298 K.

| Peptide | Run | $\alpha$-helix | Turn | Random coil |
|---|---|---|---|---|
| 1-55 | Bulk | 56.1 % | 17.4 % | 24.9 % |
| | 1 | 58.7 % | 20.2 % | 18.3 % |
| | 2 | 59.7 % | 19.7 % | 18.6 % |
| | 3 | 50.1 % | 25.1 % | 21.7 % |
| 56-131 | Bulk | 63.6 % | 14.5 % | 21.3 % |
| | 1 | 50.4 % | 22.1 % | 24.6 % |
| | 2 | 54.0 % | 20.6 % | 22.3 % |
| | 3 | 53.7 % | 20.8 % | 23.8 % |

Increasing the temperature (to 350 K) changes the structures of the two peptides in different ways (Fig. 5). Starting from the end of the first simulation run for peptide1-55 a change in the tertiary structure can be seen, with the peptide adopting a flat structure as seen for the second and third simulations. This suggests that the two different states seen in the 298 K simulations are separated by an energy barrier, which may be surmounted at higher temperatures. Similar to the 298 K runs the peptide secondary structure remains largely unchanged relative to the solution structure. For peptide56-131 (also starting from the end of the first simulation) the radius of gyration and its parallel and perpendicular components remain largely unchanged but it shows a marked decrease in ordered secondary structure. One of the interfacially adsorbed helices (residues 82-97) largely disappears across the simulation. The 104-116 helix also gets smaller and moves closer to the interface, with the peptide overall exhibiting a flatter structure.

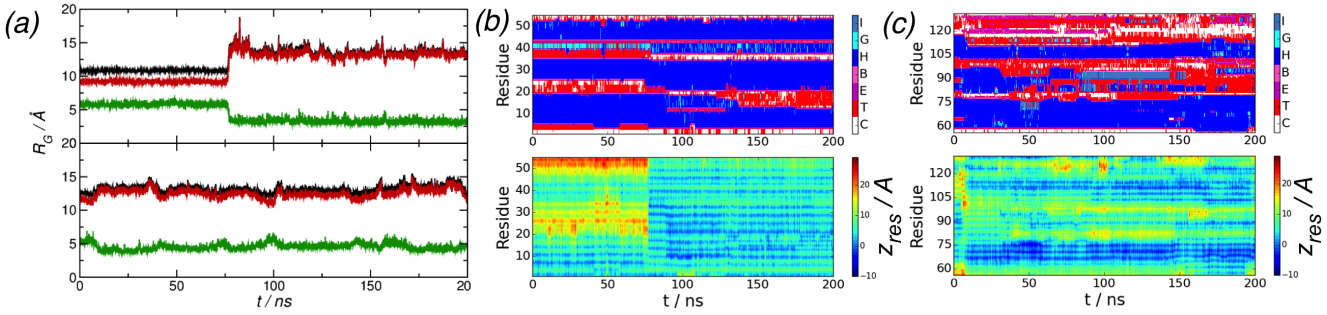



Figure 5: Peptide structure and positions at air-water interface at 350 K (a) Radius of gyration for peptide1-55 (top) and peptide56-131 (bottom). Black line denotes $R_G$, red $R_G^{xy}$, and green $R_g^z$. (b) Secondary structure (top) and residue-interface separation (bottom) for peptide1-55 at 350 K. For secondary structure blue denotes $\alpha$-helix, red turn, cyan 3-10-helix, and white random coil. (c) Secondary structure (top) and residue-interface separation (bottom) for peptide56-131 at 350 K. Symbols as in (b).

## Both peptide1-55 and 56-131 adopt multiple conformations at air-water interface

To assess the coexistence of the different interfacial conformations of peptide1-55 REST simulations[37,38] were performed. From the radius of gyration the existence of three different interfacial conformations can be seen (Fig. 6(a)), with frequent transitions between these (all discussion in this section refers to the $\lambda = 1$ as this is the only physically relevant replica). These frequent transitions can also be seen in the supporting movie peptide1-55_air-water_REST.avi. The two-dimensional probability distribution $P(R_G^{xy}, R_G^z)$ shows that the conformations seen correspond to a compact state, with similar size to the solution structure, and to two extended (extended 1 and 2) conformations. The two extended conformations are differentiated by $R_G^{xy}$, with the second, less common state (extended 2) being more extended in the plane of the interface (Fig. 6(b)). For the two extended states $R_G^z$ are essentially identical, similar to the value for a single $\alpha$-helix suggesting that these correspond to different arrangements of the helices at the air-water interface.

Comparison of representative structures for these different states shows that the compact state is similar to the solution and run1 conformations. The two extended conformations

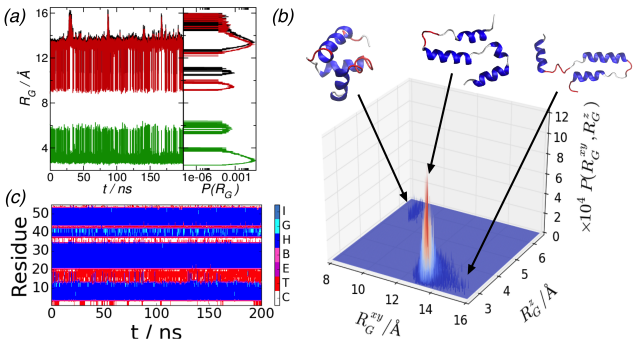



Figure 6: Structure of peptide1-55 from REST simulations. (a) (left) Radii of gyration. Black line denotes $R_G$, red line $R_G^{xy}$ and green line $R_G^z$. Right hand panel shows probability distributions for $R_G$, $R_G^{xy}$, and $R_G^z$ calculated over last 100 ns of simulation. (b) Two-dimensional probability distribution. Insets show representative structures (found using the gromacs g_cluster utility). (c) Secondary structure composition. Video file in supplementary information named peptide1-55_air-water_REST.avi shows REST simulation trajectory.

have similar structures with the two main helices lying parallel to each other. These differ by the presence of a short helical segment (38-42) present in the first extended state and absent in the second and by the location of the final helix (44-53). In the more common state this also lies parallel to the other helices, which maximises its non-bonded interactions with the rest of the peptide. In the second extended state it lies at an angle to the main body of the peptide; this increases the surface area occupied by the peptide but at the cost of decreasing favourable non-bonded interactions between this helix and the rest of the peptide. The secondary structure of the three different conformations is similar, the first extended having an additional short segment of $\alpha$-helix (residues 38-42) that is absent in the other conformations, but which is present in the crystal structure. As well as the $\alpha$-helix content increasing the distribution of these changes relative to the solution state. Compared to the solution state the first helix is somewhat shorter, while the $\alpha$-helices near the C-terminal end grow. While this end of the peptide is overall more hydrophilic than the N-terminus end the growth in helices here allows the hydrophobic residues in this region to partition into the air.

Examining the probability distributions of $R_G$ and its parallel and perpendicular com-

Table 2: Decomposition of energies found from REST simulations of peptide1-55 and peptide56-131 at air-water interface. Energies in kcal mol$^{-1}$ and surface area in Å$^2$. Uncertainties quoted are taken to be one standard deviation of the mean.

| | Peptide1-55 | | | Peptide56-131 | | |
|---|---|---|---|---|---|---|
| | Compact | Extended1 | Extended2 | Compact1 | Compact 2 | Extended |
| Count | 142 | 9819 | 41 | 9012 | 971 | 19 |
| $E_{vdw}$ | -316.2±0.8 | -282.9±0.1 | -279.4±1.7 | -413.1±0.14 | -392.7±0.4 | -373±3 |
| $E_{elec}$ | 1512±2.3 | 1523.4±0.3 | 1546±3.6 | 2152.6±0.3 | 2167.7±0.9 | 2190±10 |
| $\Delta F_{partition}$ | 9.2±0.8 | -5.2±0.1 | -1.0±2.0 | 7.2±0.11 | -5.2±0.4 | 7±5 |
| $\Delta F_{interface}$ | -12.4±0.3 | -26.3±0.1 | -27.9±0.7 | -19.87±0.03 | -24.2±0.15 | -34±1.1 |
| Surface area | 157.8±3.8 | 335.0±0.6 | 354.4±9.3 | 252.9±0.4 | 308±1.9 | 430±15 |

ponents shows that the peptide predominately adopts one extended state. The free energy difference between the compact and extended states, found using $F(R_G) = -k_{\mathrm{B}}T \ln P(R_G)$ (see Fig. S3), is $\sim$ 1.87 kcal mol$^{-1}$ ($\sim$ 3 $k_{\mathrm{B}}T$), showing that there is only a slight difference in energies between these. In a similar way the barrier between the compact and extended states can be estimated as 20.99 kcal mol$^{-1}$ (approximately 35 $k_{\mathrm{B}}T$), although due to the poor sampling of intermediate values of $R_G$ this is only an estimate. The different states can be partitioned based on $R_G$; the compact state has $R_G <$ 12 Å, the first (more populous) extended state has 12 Å$< R_G <$15 Å, while the second extended state has $R_G >$ 15 Å. Calculating the intramolecular non-bonded (van der Waals and electrostatic) energies (Table 2) and the changes in the partition and interfacial free energies ($\Delta F_{partition}$ and $\Delta F_{interface} = -\gamma A_{occupied}$ where $\gamma$ = 54.6 mN m$^{-1}$ is the air-water surface tension for SPC water[58]) can help to identify the different contributions to the stability of the different conformations. The compact structure has the lowest interaction energies of the different conformations due to its more compact structure. The interaction energy for the second extended state is slightly higher than the first, reflecting the reduction in the attractive interactions due to the placement of the 44-53 helix. The extended states are stabilised by the partition and interfacial free energies, which are both substantially lower for the extended states than for the compact one. It should be noted that the calculated surface area ($\sim$ 335 Å$^2$) is close to the value found experimentally (300 Å$^2$) by Holt *et al.*[34]

REST simulation of peptide56-131 confirms that the structural change at the the air-water interface occurs over a longer timescale than for peptide1-55 (Fig. 7(a)). The histogram of $R_G$ has three different regions; a peak at about $R_G \sim 12.2$ Å, which corresponds to the compact solution structure, a smaller peak centred around $R_G \sim 13$ Å, and some highly extended conformations with $R_G > 14$ Å. Similar peaks are seen in the $R_G^{xy}$ and $R_G^z$ histograms. The two-dimensional probability distribution (Fig. 7(b)) shows that the slightly extended state is extended in the perpendicular direction, suggesting that it occupies a slightly smaller interfacial area while penetrating further into the water. The most extended states lies flat at the interface. Unlike peptide1-55 significant differences in the secondary structure are seen between these different states, with regions of $\pi$-helix appearing in the compact 2 state and the extended state exhibiting a significant decrease in the $\alpha$-helix content. Also unlike peptide1-55 the compact structure, closest to the solution structure, is the most commonly found.

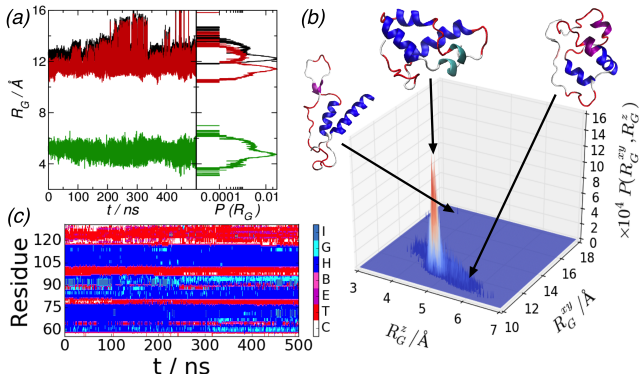



Figure 7: Structure of peptide56-131 from REST simulation. (a) (left) Radii of gyration. Black line denotes $R_G$, red line $R_G^{xy}$ and green line $R_G^z$. Right hand panel shows probability distributions for $R_G$, $R_G^{xy}$, and $R_G^z$ calculated over last 100 ns of simulation. (b) Two-dimensional probability distribution. Insets show representative structures (found using the gromacs g_cluster utility). (c) Secondary structure composition. Video file in redsupplementary information named peptide56-131_air-water_REST.avi shows REST simulation trajectory.

The REST simulations continue the trend of decreasing $\alpha$-helix content seen in the NVT simulations with the $\alpha$-helix content being $\sim$20 % lower than in bulk solution (Table 3).

The slow decrease in the $\alpha$-helix may be responsible for the longer timescales over which the structural change occurs, compared to peptide1-55. The more extended states have even lower $\alpha$-helix contents ($\sim$10-15 % lower than the compact state), which may also be seen from the representative snapshots (Fig 7(b)). As for peptide1-55 the compact state has more favourable interaction energy than the extended state, due to the close contact between the different helices. The partition free energy is lowest for the the compact 2 state; the desire to lower the partition free energy may drive the change from $\alpha$ to $\pi$-helix seen in this state. Reflecting their larger sizes the compact 2 and extended states have higher interfacial areas, while the most compact state has low intramolecular interaction energies (Table 2). Using $P(R_G)$ the free energy difference between the two compact states $\sim$ 1.85 kcal mol$^{-1}$ (Figure S4), but unlike peptide1-55 there is only a small free energy barrier between these two states, which is consistent with the higher probability of finding the peptide in this state. The more extended states are found only rarely in the simulation (0.1 % of conformations). Similar to the most extended state for peptide1-55 while these occupy a significantly higher interfacial area these have significantly weaker intramolecular interactions.

Table 3: Secondary structure composition of peptide1-55 and peptide56-131 at air-water interface from REST simulations. *Peptide56-131 compact 2 state contains 5.6 % $\pi$-helix.

| Peptide | Conformation | $\alpha$-helix | Turn | Random coil |
|---|---|---|---|---|
| 1-55 | Total | 66.2 % | 12.8 % | 20.2 % |
| | Compact | 66.2 % | 13.4 % | 20.4 % |
| | Extended 1 | 66.2 % | 12.8 % | 20.2 % |
| | Extended 2 | 63.0 % | 19.4 % | 17.6 % |
| 56-131 | Total | 44.0 % | 29.8 % | 22.5 % |
| | Compact 1 | 45.3 % | 29.8 % | 22.0 % |
| | Compact 2* | 31.9 % | 29.8 % | 26.9 % |
| | Extended | 34.9 % | 36.2 % | 25.0 % |

# Discussion and conclusions

The conformation of proteins at liquid interfaces determines their use in a number of industrial and biotechnological applications, as well as a number of biological processes. The difficulty in determining protein structure in these environments, however, has limited our understanding of protein function at interfaces. As the complex tertiary structures of proteins can make elucidating the relationship between biomolecular structure and interfacial function it is common to study simpler peptides, either *de novo* designed or fragments from larger proteins. In this paper the behaviour of two such peptides, derived from myoglobin, are studied at the air-water interface using molecular dynamics simulations.

While these two peptides share some structural similarities, having similar proportions of hydrophobic residues and forming compact, $\alpha$-helical structures in solution these exhibit very different behaviour at the air-water interface. Peptide1-55, which was found to be an effective emulsion and foam stabiliser, was found to adopt an extended structure at the interface. This maximised both the interfacial area it occupies and the partitioning of hydrophobic residues into air. Using REST the coexistence of both extended and compact conformations was found. These were found to be close in energy, with the free energy difference between these $\sim$3 $k_{\mathrm{B}}T$, but the barrier separating these states is likely to be significantly higher. Different effects are found to stabilise the different states with the compact state being stabilised by favourable intramolecular interactions, while the extended state is favoured by interfacial and partition free energies. The transition between compact and extended forms at the air-water interface is similar to that seen for the biosurfactant protein Rsn-2.[3] This peptide also has an increase in $\alpha$-helix content at the interface. The second peptide, which is a poorer emulsifier, was found to remain close to its solution structure. While REST simulation found an extended conformation this was significantly less common and in this case there was a decrease in the $\alpha$-helix content, with the central helix in particular becoming smaller. This is consistent with previous computational work which examined the isolated myoglobin helices in solution, which found that that helix was easily destabilised.[57]

In the full protein additional non-bonded interactions between the peptides studied in this work exist, so the conformations of these regions in the full protein may differ from the free peptides. Nevertheless neutron reflectivity[34] has shown that myoglobin and apo-myoglobin form layers at the air-water interface of similar thickness to peptide1-55 ($\sim 12 - 16$ Å), suggesting that at least parts of the full proteins undergo similar structural rearrangements. Interfacial conformation can also be affected by interactions between proteins, in particular in denser monolayers, which is beyond the scope of this work. This can lead to the formation of protein fibrils at liquid interfaces and other higher order structures. The conformation of individual peptides is of interest as this gives insight into the initial stages of interfacial adsorption, in particular conformational changes, which plays a role in determining their further aggregation and assembly.

The differences in the conformations of these peptides at the air-water interface allows for the development of a model for their different emulsification behaviour. By forming a more extended state peptide1-55 occupies a larger interfacial area per peptide ($\sim 335$ Å$^2$ compared to $\sim 250$ Å$^2$) so fewer molecules would be necessary to cover a given interfacial area, leading to a higher emulsification activity. This was observed in the experimental studies of peptide1-55 and 56-131, which showed that the the mean droplet diameter for peptide1-55 stabilised emulsions was substantially smaller than for peptide56-131 (4.8$\mu$m against 7.6$\mu$m), so the total interfacial area in the peptide1-55 stabilised emulsion is 58 % larger. Crucially, this increase in the area stabilised by the peptides was too large to be explained purely by the larger number of peptide1-55 molecules, suggesting a higher occupied area per peptide. The difference in the interfacial conformations of the peptides may also lead to the formation of different monolayer structures at interfaces, illustrated schematically in Fig. 8. The extended conformation of peptide1-55 allows for close approach and alignment of $\alpha$-helices on neighbouring peptides, leading to a network stabilised by strong, intermolecular interactions. As the link between the second and third $\alpha$-helices is flexible, as demonstrated by the existence of different extended states would allow for interdigitation of these, leading

to a further increase in the strength of the interfacial film. By contrast the compact form exhibited by peptide56-131 at the interface leads to weaker interactions between neighbouring molecules, decreasing the strength of the interfacial film. It should of course be noted that this assumes no further change in the interfacial conformation due to peptide aggregation and future work will investigate the aggregation of these peptides at liquid interfaces to directly test this.

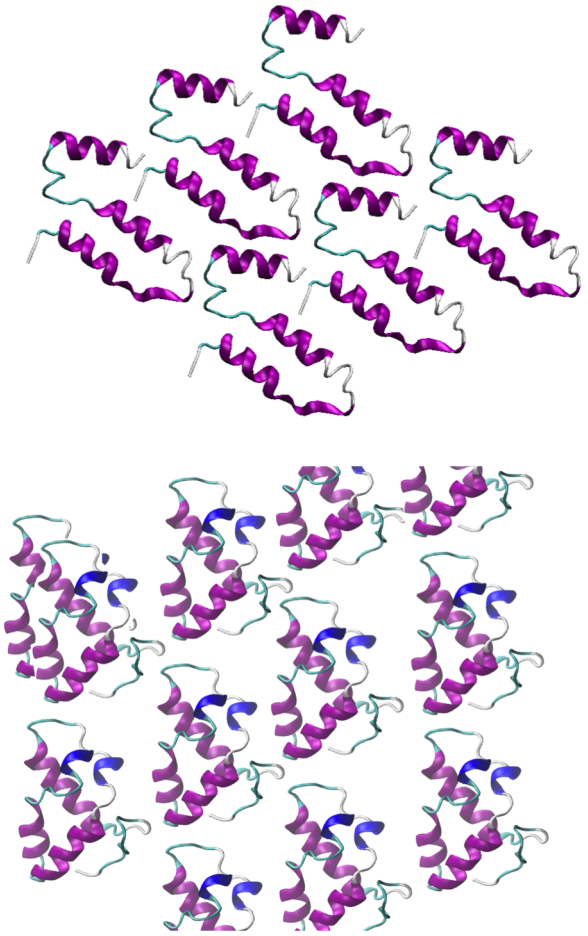


Figure 8: Schematic illustration of possible monolayer structures for peptide1-55 (top) and peptide56-131 (bottom). Peptide conformations taken to be the representative conformations for the most populated state found from the REST simulations.

Comparison between these peptides allows us to test some observations regarding the interfacial behaviour of proteins. Studies of simple model peptides have often shown that

adsorption at interfaces leads to the formation of $\alpha$-helices in otherwise disordered peptides.[17] While SRCD measurements show that some proteins, including $\alpha$-lactoalbumin[59] and $\beta$-lactoglobin,[60] also increase their $\alpha$-helix content at liquid interfaces this is not universal as others, including lysozyme,[13] show a decrease. SRCD measurements of myoglobin[13] showed that the $\alpha$-helix content deceases at oil-water interfaces, although circular dichroism measurements of myoglobin at the air-water interface show that the change in secondary structure there is pH dependent.[33] The differing changes in $\alpha$-helix content seen in the two peptides studied here further demonstrates that the conformational changes at interfaces are dependent on the protein sequence. For peptide1-55 creation of additional $\alpha$-helix allows for the hydrophobic residues in the more hydrophilic C-terminus end to partition into the air, increasing the overall proportion of $\alpha$-helix. By contrast the $\alpha$-helical content of peptide56-131 decreases upon absorption at the air-water interface with the helices in contact with the interface becoming smaller. For peptide1-55 the presence of different conformations with similar secondary structure suggests that it may adopt a 'molten globule'-like state at the air-water interface, while peptide56-131 does not show such a state. Previously such a state was suggested for $\alpha$-lactalbumin leading to improved interfacial activity.[61,62] In the case of $\alpha$-lactalbumin SRCD measurements ruled this out,[59] the present work suggests that the existence of the molten-globule state is possible for particular proteins.

As with all molecular dynamics studies the choice of force field may influence the conformations found. The Gromos family of force fields,[43,63,64] as used in this work, has been used in a large number of previous studies of proteins at liquid interfaces, often giving results that are in good agreement with experimental studies. This has been parameterised against hydration and solvation free energies of amino acid sidechain analogues and so would be expected to give a good description of the partitioning of residues between water and air , one of the main drivers for interfacial adsorption and conformational change. Nevertheless the complex nature of the makes the accurate modelling of molecular structure in this environment challenging and future work will investigate the performance of polarizable water

models on the conformations of peptides at interfaces. This emphasises the importance of quantitively validating the results of simulation work through experimental measurements. Changes in protein secondary structure at liquid interfaces may be investigated using circular dichroism, in particular SRCD and RIME CD, and FTIR, which would provide a direct test of the results in this paper. Due to the presence of a number of aromatic amino acids in the studied peptides, fluorescence spectroscopy[65] may also be used to investigate the partitioning of these residues between the water and air.

## Acknowledgements

The author wishes to acknowledge the SFI/HEA Irish Centre for High-End Computing (ICHEC) for the provision of computational facilities and support. The modified version of PLUMED for performing REST simulations was provided by Giovanni Bussi (SISSA).

## Supporting information available

Populations of conformations and acceptance probabilities for peptide1-55 and peptide56-131 from REST simulations, gyration tensor eigenvalues for peptide56-131 at the air-water interface, and free energies for peptide1-55 and 56-131 from REST simulations. PDB files containing coordinates of representative structures from Figs. 6 and 7. Videos of REST simulation trajectories (peptide1-55_air-water_REST.avi and peptide56-131_air-water_REST.avi).

### References

(1) Yano, Y. F. Kinetics of protein unfolding at interfaces. *J. Phys. Condens. Matter* **2012**, *24*, 503101.

(2) Linder, M. B. Hydrophobins: Proteins that self assemble at interfaces. *Curr. Opin. Colloid Interface Sci.* **2009**, *14*, 356–363.

(3) Mackenzie, C. D.; Smith, B. O.; Meister, A.; Blume, A.; Zhao, X.; Lu, J. R.; Kennedy, M. W.; Cooper, A. Ranaspumin-2: structure and function of a surfactant protein from the foam nests of a tropical frog. *Biophys. J.* **2009**, *96*, 4984–92.

(4) Vance, S. J.; Mcdonald, R. E.; Cooper, A.; Smith, B. O.; Kennedy, M. W. The structure of latherin, a surfactant allergen protein from horse sweat and saliva. *J. R. Soc. Interface* **2013**, *10*, 20130453.

(5) Reis, P.; Holmberg, K.; Watzke, H.; Leser, M. E.; Miller, R. Lipases at interfaces: a review. *Adv. Colloid Interface Sci.* **2008**, *147-148*, 237–50.

(6) Dickinson, E. Colloid science of mixed ingredients. *Soft Matter* **2006**, *2*, 642.

(7) DeGrado, W.; Lear, J. Induction of peptide conformation at apolar water interfaces. 1. A study with model peptides of defined hydrophobic periodicity. *J. Am. Chem. Soc.* **1985**, *107*, 7684–7689.

(8) Dexter, A. F.; Malcolm, A. S.; Middelberg, A. P. J. Reversible active switching of the mechanical properties of a peptide film at a fluid-fluid interface. *Nat. Mater.* **2006**, *5*, 502–6.

(9) Gidalevitz, D.; Huang, Z.; Rice, S. A. Protein folding at the air-water interface studied with x-ray reflectivity. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 2608–11.

(10) Zhao, X.; Pan, F.; Lu, J. R. Interfacial assembly of proteins and peptides: recent examples studied by neutron reflection. *J. R. Soc. Interface* **2009**, *6*, S659–70.

(11) Husband, F. A.; Garrood, M. J.; Mackie, A. R.; Burnett, G. R.; Wilde, P. J. Adsorbed protein secondary and tertiary structures by circular dichroism and infrared spectroscopy with refractive index matched emulsions. *J. Agric. Food Chem.* **2001**, *49*, 859–866.

(12) Miles, A. J.; Wallace, B. A. Synchrotron radiation circular dichroism spectroscopy of proteins and applications in structural and functional genomics. *Chem. Soc. Rev.* **2006**, *35*, 39–51.

(13) Zhai, J. L.; Day, L.; Aguilar, M.-I.; Wooster, T. J. Protein folding at emulsion oil/water interfaces. *Curr. Opin. Colloid Interface Sci.* **2013**, *18*, 257–271.

(14) Day, L.; Zhai, J.; Xu, M.; Jones, N. C.; Hoffmann, S. V.; Wooster, T. J. Conformational changes of globular proteins adsorbed at oil-in-water emulsion interfaces examined by Synchrotron Radiation Circular Dichroism. *Food Hydrocoll.* **2014**, *34*, 78–87.

(15) Euston, S. R.; Hughes, P.; Naser, M. A.; Westacott, R. E. Comparison of the adsorbed conformation of barley lipid transfer protein at the decane-water and vacuum-water interface: a molecular dynamics simulation. *Biomacromolecules* **2008**, *9*, 1443–53.

(16) Euston, S. R. Molecular simulation of adsorption of hydrophobin HFBI to the air-water, DPPC-water and decane-water interfaces. *Food Hydrocoll.* **2013**, *42*, 66–74.

(17) Dalgicdir, C.; Globisch, C.; Peter, C.; Sayar, M. Tipping the Scale from Disorder to Alpha-helix: Folding of Amphiphilic Peptides in the Presence of Macroscopic and Molecular Interfaces. *PLOS Comput. Biol.* **2015**, *11*, e1004328.

(18) Zare, D.; McGrath, K. M.; Allison, J. R. Deciphering $\beta$-Lactoglobulin Interactions at an Oil-Water Interface: A Molecular Dynamics Study. *Biomacromolecules* **2015**, *16*, 1855–1861.

(19) Euston, S. R. Molecular dynamics simulation of protein adsorption at fluid interfaces: a comparison of all-atom and coarse-grained models. *Biomacromolecules* **2010**, *11*, 2781–7.

(20) Cheung, D. L. Molecular simulation of hydrophobin adsorption at an oil-water interface. *Langmuir* **2012**, *28*, 8730–6.

(21) Schulz, A.; Fioroni, M.; Linder, M. B.; Nessel, A.; Bocola, M.; Subkowski, T.; Schwaneberg, U.; Böker, A.; Rodríguez-Ropero, F. Exploring the mineralization of hydrophobins at a liquid interface. *Soft Matter* **2012**, *8*, 11343–11352.

(22) James, J. J.; Lakshmi, B. S.; Seshasayee, A. S. N.; Gautam, P. Activation of Candida rugosa lipase at alkane-aqueous interfaces: a molecular dynamics study. *FEBS Lett.* **2007**, *581*, 4377–83.

(23) Xue, Y.; He, L.; Middelberg, A. P. J.; Mark, A. E.; Poger, D. Determining the Structure of Interfacial Peptide Films: Comparing Neutron Reflectometry and Molecular Dynamics Simulations. *Langmuir* **2014**,

(24) Iglesias-Fernández, J.; Darré, L.; Kohlmeyer, A.; Thomas, R. K.; Shen, H.-H.; Domene, C. Surfactin at the Water/Air Interface and in Solution. *Langmuir* **2015**, *31*, 11097–11104.

(25) Trodler, P.; Schmid, R. D.; Pleiss, J. Modeling of solvent-dependent conformational transitions in Burkholderia cepacia lipase. *BMC Struct. Biol.* **2009**, *9*, 38.

(26) Sugita, Y.; Okamoto, Y. Replica exchange molecular dynamics method for protein folding simulation. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(27) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *The Journal of Chemical Physics* **2004**, *120*, 11919–11929.

(28) De Simone, A.; Kitchen, C.; Kwan, A. H.; Sunde, M.; Dobson, C. M.; Frenkel, D. Intrinsic disorder modulates protein self-assembly and aggregation. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6951–6956.

(29) Deighan, M.; Pfaendtner, J. Exhaustively Sampling Peptide Adsorption with Metadynamics. *Langmuir* **2013**, *29*, 7999–8009.

(30) Levine, Z. A.; Fischer, S. A.; Shea, J. E.; Pfaendtner, J. Trp-Cage Folding on Organic Surfaces. *J. Phys. Chem. B* **2015**, *119*, 10417–10425.

(31) Watson, H. C. The Stereochemistry of the Protein Myoglobin. *Prog. Stereochem.* **1969**, *4*, 299.

(32) Poon, S.; Clarke, A. E.; Schultz, C. J. Structure-function analysis of the emulsifying and interfacial properties of apomyoglobin and derived peptides. *J. Colloid Interface Sci.* **1999**, *213*, 193–203.

(33) Sankaranarayanan, K.; Dhathathreyan, A.; Krägel, J.; Miller, R. Interfacial viscoelasticity of myoglobin at air/water and air/solution interfaces: Role of folding and clustering. *J. Phys. Chem. B* **2012**, *116*, 895–902.

(34) Holt, S. A.; McGillivray, D. J.; Poon, S.; White, J. W. Protein Deformation and Surfactancy at an Interface. *J. Phys. Chem. B* **2000**, *104*, 7431–7438.

(35) Kyte, J.; Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105–132.

(36) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.; Appel, R.; A., B. *Proteomics Protoc. Handb.*; 2005; pp 571–601.

(37) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. Replica exchange with solute tempering: a method for sampling biological systems in explicit water. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13749–54.

(38) Bussi, G. Hamiltonian replica exchange in GROMACS: a flexible implementation. *Mol. Phys.* **2013**, *112*, 379–384.

(39) Abrams, C.; Bussi, G. Enhanced Sampling in Molecular Dynamics Using Metadynamics, Replica-Exchange, and Temperature-Acceleration. *Entropy* **2014**, *16*, 163–199.

(40) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(41) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(42) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604–613.

(43) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; Van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843–856.

(44) Berendsen, H.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(45) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577.

(46) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 1–7.

(47) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

(48) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of MolecularDynamics Simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327.

(49) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; Van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chemie Int. Ed.* **1999**, *38*, 236–240.

(50) Vink, R. L. C.; Horbach, J.; Binder, K. Capillary waves in a colloid-polymer interface. *J. Chem. Phys.* **2005**, *122*, 134905.

(51) Radzicka, A.; Wolfenden, R. Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. *Biochemistry* **1988**, *27*, 1664–1670.

(52) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(53) Frishman, D.; Argos, P. Knowledge-based protein secondary structure assignment. *Proteins Struct. Funct. Genet.* **1995**, *23*, 566–579.

(54) Park, S.; Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **2004**, *120*, 5946–61.

(55) Fiorin, G.; Klein, M. L.; Hénin, J. Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* **2013**, *111*, 3345–3362.

(56) Laio, A.; Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports Prog. Phys.* **2008**, *71*, 126601.

(57) Hirst, J. D.; Brooks, C. L. Molecular Dynamics Simulations of Isolated Helixes of Myoglobin. *Biochemistry* **1995**, *34*, 7614–7621.

(58) Vega, C.; Miguel, E. D. Surface tension of the most popular models of water by using the test-area simulation method. *J. Chem. Phys.* **2007**, *126*, 154707.

(59) Zhai, J.; Hoffmann, S. V.; Day, L.; Lee, T.-H.; Augustin, M. A.; Aguilar, M.-I.; Wooster, T. J. Conformational changes of $\alpha$-lactalbumin adsorbed at oil-water interfaces: interplay between protein structure and emulsion stability. *Langmuir* **2012**, *28*, 2357–67.

(60) Zhai, J.; Miles, A. J.; Pattenden, L. K.; Lee, T.-H.; Augustin, M. A.; Wallace, B. A.; Aguilar, M.-I.; Wooster, T. J. Changes in beta-lactoglobulin conformation at the oil/water interface of emulsions studied by synchrotron radiation circular dichroism spectroscopy. *Biomacromolecules* **2010**, *11*, 2136–42.

(61) Wijesinha-Bettoni, R.; Gao, C.; Jenkins, J. A.; Mackie, A. R.; Wilde, P. J.; Mills, E. N. C.; Smith, L. J. Heat treatment of bovine $\alpha$-lactalbumin results in partially folded, disulfide bond shuffled states with enhanced surface activity. *Biochemistry* **2007**, *46*, 9774–9784.

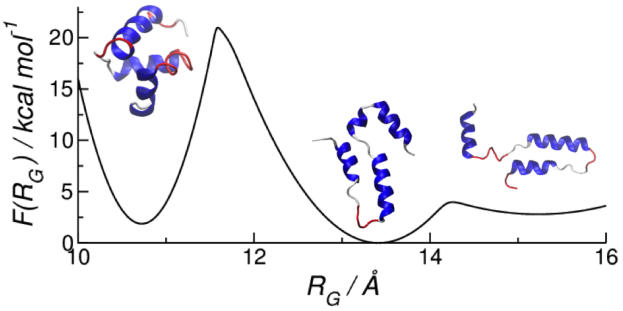
(62) Gao, C.; Wijesinha-Bettoni, R.; Wilde, P. J.; Mills, E. N. C.; Smith, L. J.; Mackie, A. R. Surface Properties Are Highly Sensitive to Small pH Induced Changes in the 3-D Structure of $\alpha$-Lactalbumin. *Biochemistry* **2008**, *47*, 1659–1666.

(63) van Gunsteren, W. F.; Daura, X.; Mark, A. E. *Encyclopedia of Computational Chemistry*; John Wiley & Sons, Ltd, 2002.

(64) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(65) Sakakibara, K.; Fujisawa, T.; Hill, J. P.; Ariga, K. Conformational interchange of a carbohydrate by mechanical compression at the air-water interface. *Phys. Chem. Chem. Phys.* **2014**, *16*, 10286–94.

Table of contents graphic



# Supporting Information for 'Conformations of myoglobin-derived peptides at the air-water interface'

David L. Cheung

*School of Chemistry, National University of Ireland Galway, Galway, Ireland**

## I. CONVERGENCE OF REST SIMULATIONS

Shown in Fig. S1 is a plot showing the relative populations of the main conformational states of peptide1-55 and peptide56-131 throughout the REST simulations. Specifically for every hundred saved datasets the population of the three conformation states, identified from Figs. 5 and 6, were calculated. For peptide 1-55, apart from a very brief period at the beginning of the simulation the populations of compact and extended 1 states were $\sim 0.015$ and $\sim 0.98$ throughout the simulation. The population of the extended 1 state briefly decreases when the peptide intermittently adopts the most extended conformation (around 30 ns, 90 ns, and 170 ns).

For peptide 56-131 the populations take longer to reach constant values (approximately 300 ns). The populations of the two main states remain approximately 0.9 and 0.1 across the last 200 ns of the simulation with brief, intermittent periods where they become close to equally populated (about 350 ns and 450 ns). This consistent with the lower free energy barrier between these states. In both cases no systematic change was observed in the populations across the last 150 ns of the simulations.

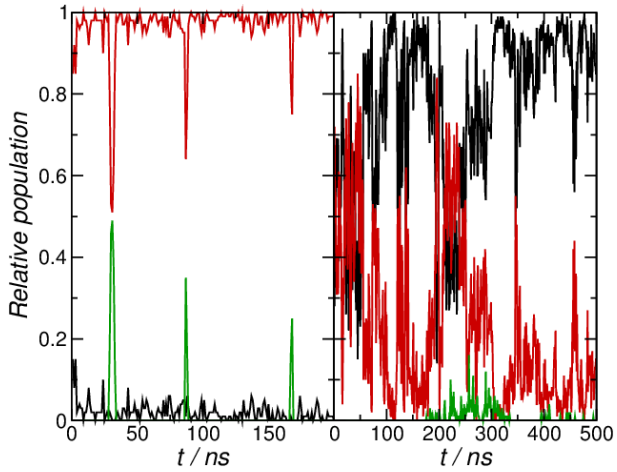



Figure S1 Relative populations of peptide conformations from the REST simulations of peptide1-55 (left) and peptide56-131 (right), calculated from block averaging over 100 consecutive data sets. For peptide1-55 black line denotes compact conformation, red line extended 1, and green line extended 2 and for peptide56-131 black line denotes compact 1 conformation, red line extended 1, and green line extended 2 (see text for discussion of the different conformations).

## II. ACCEPTANCE PROBABILITIES FOR REST SIMULATIONS

Table S1 lists the acceptance probabilities for replica exchanges in the REST simulations. For both peptides these were typically 20 %-30 %, within the typical range for replica exchange simulations.

| Peptide | 298 K↔307.4 K | 307.4 K↔317.8 K | 317.8 K↔328.9 K | 328.9 K↔338.92 K | 338.92 K↔350 K |
|---|---|---|---|---|---|
| Peptide1-55 | 31.9 % | 31.3 % | 24.3 % | 29.2 % | 31.0 % |
| Peptde56-131 | 30.7 % | 30.2 % | 30.8 % | 33.7 % | 35.3 % |

Table S1 Acceptance probabilities for REST simulations of peptide1-55 and 56-131 at air-water interface.

## III. GYRATION TENSOR EIGENVALUES FOR PEPTIDE56-131 AT THE AIR-WATER INTERFACE

Shown in Fig. S2 are the gyration tensor eigenvalues for peptide56-131 from NVT simulations at the air-water interface. The values of these are similar for all three replicate runs and are similar to the values from bulk simulation. This indicates that this peptide exhibits only a single, globular conformation in the NVT simulations, suggesting that the difference in $R_G^{xy}$ and $R_G^z$ arises due to changes in peptide orientation rather peptide conformation.

## IV. PROBABILITY DISTRIBUTION AND FREE ENERGY FOR PEPTIDE1-55

The free energy differences between the different states can be estimated from the probability distribution using

$$F\left(R_G\right) = -k_{\mathrm{B}} T \log P\left(R_G\right) \tag{1}$$

To aid this $P(R_G)$ is approximated by the sum of three Gaussians

$$P_{fit}(R_G) = a_1 \exp\left(-b_1(R_G - c_1)^2\right) + a_2 \exp\left(-b_2(R_G - c_2)^2\right) + a_3 \exp\left(-b_3(R_G - c_3)^2\right) \tag{2}$$

with the coefficients listed in Table S2. From Fig. S3 this is a good approximation to the simulation probability distribution around the two main peaks in $P(R_G)$. $F\left(R_G\right)$ can then

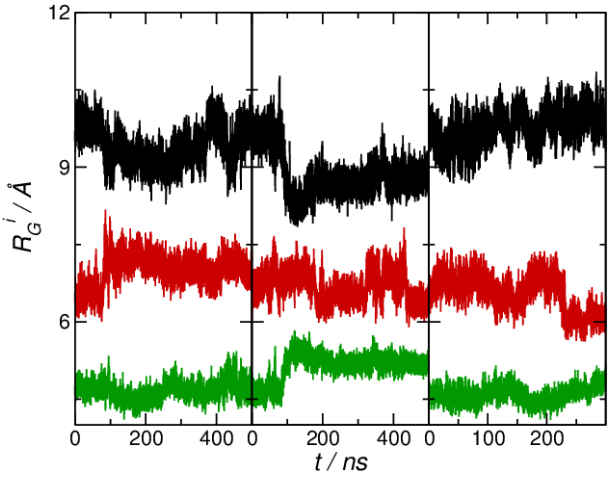



Figure S2 Gyration tensor eigenvalues for peptide56-131 at air-water interface at 298 K for run 1, 2, and 3 (left to right). Black denotes $R_G^1$, red $R_G^2$, and green $R_G^3$.

be used to estimate the free energy difference between the compact and extended 1 states ($\sim$ 1.87 kcal mol$^{-1}$) and the barrier separating these ($\sim$ 20.99 kcal mol$^{-1}$).

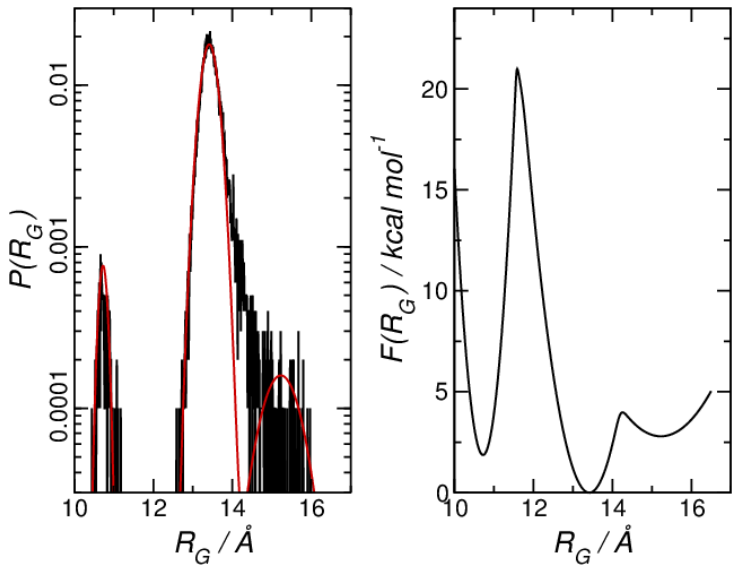



Figure S3 (left) Probability histogram $P(R_G)$ for peptide1-55 at air-water interface from REST simulation. Black line is simulation $P(R_G)$, red analytical fit (Eq. 2). (right) Free energy profile $F(R_G)$.

| Conformation | $a_i$ | $b_i$ / Å$^{-1}$ | $c_i$ / Å |
|---|---|---|---|
| Compact | $7.63\times10^{-4}$ | 45.48 Å$^{-1}$ | 10.73 Å |
| Extended 1 | 0.0179 | 11.85 Å$^{-1}$ | 13.42 Å |
| Extended 2 | $1.60\times10^{-4}$ | 2.34 Å$^{-1}$ | 15.23 Å |

Table S2 $P(R_G)$ fitting parameters for peptide1-55.

## V. PROBABILITY DISTRIBUTION AND FREE ENERGY FOR PEPTIDE56-131

As for peptide1-55 the probability distribution $P(R_G)$ may be used to estimate the free energy difference between the two states. For peptide56-131 can be approximated by the sum of two Gaussians

$$P_{fit}(R_G) = a_1 \exp\left(-b_1(R_G - c_1)^2\right) + a_2 \exp\left(-b_2(R_G - c_2)^2\right) \tag{3}$$

with the coefficients in Table S3 (the population of the extended state is so small it is neglected in this analysis). From $F\left(R_G\right)$ the free energy difference between the two compact states is $\sim 1.85$ kcal mol$^{-1}$ and the barrier is $\sim 2.28$ kcal mol$^{-1}$.

| Conformation | $a_i$ | $b_i$ / Å$^{-1}$ | $c_i$ / Å |
|---|---|---|---|
| Compact 1 | 0.0325 | 42.67 Å$^{-1}$ | 12.22 Å |
| Compact 2 | 0.0143 | 5.49 Å$^{-1}$ | 12.99 Å |

Table S3 $P(R_G)$ fitting parameters for peptide56-131

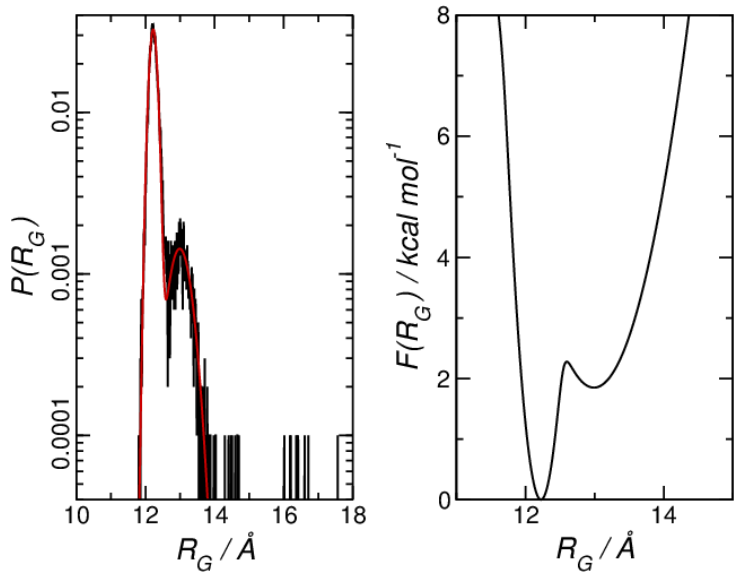



Figure S4 (left) Probability histogram $P(R_G)$ for peptide56-131 at air-water interface from REST simulation. Black line is simulation $P(R_G)$, red analytical fit (Eq. 2). (right) Free energy profile $F(R_G)$.