J Membrane Biol (2015) 248:395–405
DOI 10.1007/s00232-014-9738-9

Enhanced Sampling of Coarse-Grained Transmembrane-Peptide Structure Formation from Hydrogen-Bond Replica Exchange

Tristan Bereau · Markus Deserno

Received: 30 May 2014 / Accepted: 27 September 2014 / Published online: 14 October 2014

Abstract Protein structure formation in the membrane highlights a grand challenge of sampling in computer simulations, because kinetic traps and slow dynamics make it difficult to find the native state. Exploiting increased fluctuations at higher temperatures can help overcome free-energy barriers, provided the membrane's structure remains stable. In this work, we apply Hamiltonian replica-exchange molecular dynamics, where we only tune the backbone hydrogen-bond strength to help reduce the propensity of long-lived misfolded states. Using a recently developed coarse-grained model, we illustrate the robustness of the method by folding different WALP transmembrane helical peptides starting from stretched, unstructured conformations. We show the efficiency of the method by comparing to simulations without enhanced sampling, achieving folding in one example after significantly longer simulation times. Analysis of the bilayer structure during folding provides insight into the local membrane deformation during helix formation as a function of chain length (from 16 to 23 residues). Finally, we apply our method to fold the 50-residue-long major pVIII coat protein (fd coat) of the filamentous fd bacteriophage. Our results agree well with experimental structures and atomistic simulations based on implicit membrane models, suggesting that our explicit CG folding protocol can serve as a starting point for better-refined atomistic simulations in a multiscale framework.

T. Bereau (✉)
Max Planck Institute for Polymer Research,
Ackermannweg 10, 55128 Mainz, Germany
e-mail: bereau@mpip-mainz.mpg.de

M. Deserno
Department of Physics, Carnegie Mellon University, Pittsburgh,
PA 15213, USA

Keywords Membrane peptide · Folding · Coarse-grained · Simulations · WALP

Introduction

A thorough understanding of protein–lipid interactions is a formidable challenge encompassing many length and time scales. Though membrane-protein crystallization remains an arduous endeavor, steady progress is being made—as witnessed from the exponentially growing number of structures in the protein data bank (Bernstein et al. 1977; Caffrey 2003; DeLucas 2009; Heijne 2011).

While computer simulators need not worry about crystallization issues, they fight against two other major hurdles: force-field accuracy (Freddolino et al. 2009; Piana et al. 2011) and sampling (Grossfield et al. 2007; Freddolino et al. 2010; Neale et al. 2011; Paloncýová et al. 2012; Neale et al. 2013; Filipe et al. 2014). The latter is severely amplified when probing a protein in a crowded membrane environment—a highly ordered and viscous medium that leaves little room for a protein to move. It should, therefore, come as no surprise that simulations of membrane–peptide folding events reported in the literature have inspired a variety of protocols to help overcome forbiddingly long time scales.

The first such method consists of sampling at elevated temperatures, where one takes advantage of larger fluctuations and a more efficient sampling of phase space. In this way, Ulmschneider et al. managed to fold WALP23 (more below) in atomistic simulations (Ulmschneider et al. 2010). The higher the temperature, the larger the fluctuations, and thus the more efficient the method will be. However, increasing the temperature too much might shift the

equilibrium state away from what one would obtain at physiological conditions. Without additional constraints on the system, the temperature at which one can run the simulation is limited by the stability of both the peptide and the membrane.

In order to probe the peptide at its physiological temperature, one can turn to more sophisticated enhanced-sampling methods, such as replica exchange (Swendsen and Wang 1986). Replica-exchange algorithms swap copies of the system simulated under different conditions (e.g., temperature) according to a Metropolis–Hastings criterion, ensuring a coupling between replicas while sampling correct canonical distributions. When swapping temperatures, replica-exchange simulations effectively allow high-temperature simulations to "feed" the lower-temperature replicas, thereby accelerating the journey toward the thermodynamically stable state (Sugita and Okamoto 1999). While denaturing the protein at high temperatures enhances sampling, destabilizing the membrane structure imposes severe limitations, because the combined self-assembly of the membrane and the folding of the peptide requires too much phase-space exploration. In a reported application of WALP16 insertion and folding from all-atom replica-exchange simulations, the stability of the membrane was ensured at all temperatures by restraining the normal distance of the lipid head groups to the bilayer midplane (Nymeyer et al. 2005). While not implausible, this constraint restricts the ensemble being sampled to a potentially biased subset of the canonical state. Though the authors sampled a constant-area ensemble, replica exchange at constant pressure, which additionally includes pressure and volume in the acceptance criterion (e.g., Okabe et al. 2001; Paschek and García 2004), should be conceivable with this method. We also point out recent developments toward so-called surface-tension replica-exchange methods, where each replica samples states subject to a different lateral tension, applied to a DPPC membrane and a transmembrane WALP peptide in a POPC membrane (Mori et al. 2013).

The last method, arguably the most efficient but also most simplified, does without any explicit description of the water/membrane environments, but instead describes each as a background continuum that implicitly affects local lipid–lipid interactions. By removing all molecular friction and distinguishing the water from the membrane environment by different dielectrics, the insertion and folding of short synthetic peptides—including WALP16, WALP19, and WALP23—was achieved (Im and Brooks 2005). In addition, the folding of the longer fd coat protein (50 residues, more below) resulted in excellent agreement with structure inferred on the basis of solid-state NMR data (Im and Brooks 2004). These studies also relied on a replica-exchange algorithm.

In light of accessing longer time scales, coarse-grained (CG) models provide a second means to navigate phase space more rapidly: by averaging over degrees of freedom, CG quasi-particles or beads not only reduce the number of particles in the system, but they also remove some molecular friction. CG models have enjoyed increasing attention, due to ever-growing needs to probe long length and time scales, as well as theoretical advances toward the development of more accurate models—see Noid (2013), Ingólfsson et al. (2014) for extensive reviews on CG models for biomolecular systems. Though coarse-graining has already provided much insight into peptide–lipid interactions (e.g., Baaden and Marrink 2013)—due in no small part to the MARTINI force field (Monticelli et al. 2008)—this approach has so far contributed little to peptide structure formation in the membrane. The main factor hampering CG models' progress in this direction is the difficulty to describe secondary structure at the CG level, especially without explicitly biasing the force field (Kar and Feig 2014). As an example of a promising strategy, we point out the recent extension of the PRIMO force field to account for the presence of a membrane, described here also as a background continuum (Kar et al. 2014). Following a different approach, the PACE force field describes peptides atomistically but uses a CG solvent (Wan et al. 2011). While these two force fields have already shown that they can stabilize various membrane peptides and proteins, their folding capabilities haven't yet been documented. Alternatively, we recently presented a CG peptide-membrane cross-parametrization that can stabilize different basic secondary structural motifs (such as α-helices and β-sheets) dependent on sequence, thermodynamic conditions, and environment, combined with a systematically coarse-grained membrane that can both self assemble into a bilayer and reproduce a number of elastic properties semiquantitatively (Bereau et al. 2014). Using this model, we observed insertion and folding of WALP23 into a transmembrane helix at elevated temperatures. Compared to the aforementioned models, the peptide-membrane cross-parametrization describes peptides with less chemical detail—its ability to fold secondary structure thus stands as an interesting endeavor.

The observation that we could not get a membrane-inserted WALP23 to fold at physiological temperature (Bereau et al. 2014) lead us to question whether the quality of the CG force field or the sampling methodology was responsible for this discrepancy. In this work, we rely on Hamiltonian replica-exchange molecular dynamics (HREMD) (Bunker and Dünweg 2000) to help sample peptide conformations in the membrane. HREMD stems from a generalization of temperature-based replica-exchange (TREMD) simulations, whereby only part of the Hamiltonian gets altered, instead of a temperature change that acts on all degrees of freedom concertedly. HREMD protocols often simply scale down the strength of a number

of interaction potentials (though counterexamples exist, e.g., Lyman et al. 2006), and we point out that a uniform weakening of all degrees of freedom strictly reduces to TREMD. The coupling of all degrees of freedom in TREMD makes the number of replicas scale strongly with system size. Replica-exchange solute tempering (REST) (Liu et al. 2005) attempts to circumvent this shortcoming by removing a (possibly large) number of solvent degrees of freedom from the reweighted acceptance criterion—the quantity that determines the exchange probability. HREMD, which by construction couples only the degrees of freedom the Hamiltonian actually alters, can also benefit from wider replicas without altering the definition of the acceptance criterion. HREMD has been applied on a number of (bio)systems to enhance sampling (see, e.g., Liu et al. 2006; Affentranger et al. 2006; Mu 2009; Jiang and Roux 2010). The algorithm was most recently applied to the folding of the $(AAQAA)_3$ peptide using an atomistic polarizable force field (Huang and MacKerell 2014). Naturally, many variants of HREMD exist, because there is no unique way of biasing a Hamiltonian to enhance sampling, and thus the question "*which* degrees of freedom to modify?" becomes prominent. Singling out the right degrees of freedom can greatly help equilibrate the system of interest. In addition, it can provide the means to avoid altering other degrees of freedom—focus on the peptide and leave the membrane unaffected. Our initial ansatz will stem from past observations of long-lived misfolded states (e.g., incomplete α-helix) carrying non-native backbone hydrogen bonds (H-bonds). We propose to tune the H-bond strength so that the strongly altered replicas do not stabilize any significant secondary structure (i.e., α-helix and β-sheet, mainly).

We apply the method to study structure formation of simple peptides in a phospholipid membrane. Following Im and Brooks (Im and Brooks 2004), we avoid the issue of insertion by starting from stretched peptide configurations that span the bilayer across its normal. Though the associated insertion process holds interest (and most importantly the insertion/folding mechanism Nymeyer et al. 2005; Ulmschneider and Ulmschneider 2008), it is hampered by the energetic barrier of one capped terminus moving across the lipid membrane. Moreover, the CG model does not provide the ability to study folding pathways (e.g., "does insertion precede folding?"), given that its parametrization only targeted equilibrium properties, and thus does not ensure the reliability of its dynamics. The present study thus only focuses on the free-energy minima the model predicts. We note that we observed insertion of WALP23 at various temperatures (including physiological) in our previous work (Bereau et al. 2014). We hereby focus on the more difficult aspect—folding—and leave the previously observed insertion aspect aside. The results presented below illustrate that the folding of a pre-inserted peptide still requires significant sampling. Using HREMD, we simulate the three abovementioned WALP peptides and probe how the peptide affects the membrane during structure formation. Finally, we show that HREMD also works efficiently for the 50-residue-long fd coat protein, where the CG model shows good agreement with the all-atom simulations with implicit water/membrane (Im and Brooks 2004), as well as structures proposed on the basis of NMR measurements (Almeida and Opella 1997; Marassi and Opella 2003).

Methods

CG Models

In the following, we briefly summarize the two CG peptide (Bereau and Deserno 2009) and lipid (Wang and Deserno 2010) models, as well as their cross parametrization (Bereau et al. 2014).

Peptides are described at the CG level using an implicit-solvent description (Bereau and Deserno 2009) that includes amino-acid specificity and can stabilize different secondary structures using a single parametrization—without explicit bias toward one conformation. Each amino acid is represented by four beads, one for the side chain and three for the backbone, thereby providing the means to describe backbone dihedrals. Phenomenological interactions were incorporated to include different system properties of interest, such as hydrophobocity and backbone H-bonds. Interaction parameters were tuned to reproduce the Ramachandran plot of tripeptides and fold a *de novo* three-helix bundle. Using a unique force field, the model can fold simple helical peptides and assemble β-sheet-rich oligopeptides (Bereau and Deserno 2009). Further applications of the model include the thermodynamics of helix-folding transitions (Bereau et al. 2010, 2011) and the formation of β-barrels at the interface between virus capsid proteins (Bereau et al. 2012). Though the model lacks chemical detail to fold more complex protein structures, the use of stabilizing elastic networks (e.g., Periole et al. 2009; Globisch et al. 2013) on parts of the protein can help explore the conformational variability of flexible segments.

For the description of the membrane, we rely on Wang and Deserno's lipid force field (Wang and Deserno 2010). This implicit solvent model maps a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) lipid into 16 beads, distinguished in 8 bead types that model different chemical moieties. Iterative-Boltzmann inversion of an all-atom POPC membrane aimed at reproducing the radial distribution functions of the CG beads. An additional phenomenological attractive interaction was included to mimick the confining

effect of water—coarse-grained out in the implicit-solvent model. The model can reproduce a number of important properties of the membrane: self assembly into the bilayer state, elastic properties (e.g., bending and stretching moduli), mass density profile, and the orientation of intramolecular bonds (Wang and Deserno 2010). Furthermore, Wang and Deserno showed that other neutral lipids could be constructed from the same set of CG bead types and reached satisfying transferability in terms of structure, area per lipid, and temperature dependence of the main phase transition (Wang and Deserno 2010).

The cross-parametrization between the two models (Bereau et al. 2014) aims to reproduce atomistic potential of mean force (PMF) curves for the insertion of individual amino acid side chains into a DOPC bilayer (MacCallum et al. 2008). Beyond experimental partitioning coefficients, these PMFs provide additional spatial resolution—normal to the bilayer plane—to parametrize the cross interactions between amino-acid side-chain and lipid bead types. A number of structural properties specific to membrane peptides were investigated to validate this simple cross-parametrization, such as tilt and hydrophobic mismatch, transient pore formation from the cooperative action of antimicrobial peptides, and folding of WALP23 at elevated temperature (Bereau et al. 2014).

Simulation Details

The CG units used throughout were constructed from a length $\mathcal{L} = 1\,\text{Å}$, an energy $\mathcal{E} = k_\text{B}T_\text{body} \approx 0.617$ kcal/mol at $T_\text{body} = 310$ K, and a mass $\mathcal{M}$. As the two models were constructed from different values of $\mathcal{M}$ (see Bereau and Deserno 2009; Wang and Deserno 2010), the resulting time units, $\tau = \mathcal{L}\sqrt{\mathcal{M}/\mathcal{E}}$, differ slightly—$\tau \approx 0.1$ ps and 0.06 ps for the peptide and lipid models, respectively. This difference is further complicated by the reduction of molecular friction during coarse-graining, where the speedup gained by the model may not be the same for different dynamical processes. Case in point: the calibration of lateral lipid diffusion from the CG simulations hints at a 10^2-fold speedup (Wang and Deserno 2010), while the peptide folding kinetics of an α-helix and a β-hairpin suggest 10^3 (Bereau 2011). As such, we refrain from extracting dynamical properties observed from the model and only express timescales in the model's natural unit, τ. In any case, for the purpose of our study absolute times are less relevant than the question whether the simulation has managed to arrive at equilibrated configurations, and the latter can be addressed without any time mapping.

We ran all simulations with the ESPResSo molecular dynamics software package (Limbach et al. 2006). Simulations were done at constant temperature at $T = 1.0\,\mathcal{E}/k_\text{B}$ using a Langevin thermostat, with constant lateral tension and vertical box height (i.e., $N\Sigma L_z T$) ensembles systematically produced at $\Sigma = 0$ from a modified Andersen barostat (Kolb and Dünweg 1999) (see Wang and Deserno 2010 for details). While the lipid model allows an integration time step $\delta t = 0.1\,\tau$, the peptide model requires a smaller value, $\delta t = 0.01\,\tau$, used throughout for the entire system.

Our systems consisted of a $(100\,\text{Å})^2$ patch of 288 pre-equilibrated POPC lipids. We capped all peptides by termini, modeled as additional amino acids with specific interactions (Bereau et al. 2014). Peptide structures were generated from high-temperature simulations in an aqueous environment, ensuring the absence of any secondary structure (see, e.g., Fig. 2a), as monitored by the STRIDE assignment algorithm (Frishman and Argos 1995). Stretched peptide configurations were inserted across the membrane such that the two termini were located on opposite sides of the bilayer. All scripts necessary to prepare, run, and analyze peptide-membrane simulations are available online (http://plumo.googlecode.com).

Enhanced-Sampling Protocol

The enhanced-sampling protocol carried out here consists of a Hamiltonian replica-exchange molecular dynamics (HREMD) methodology (Bunker and Dünweg 2000), where we solely tune the strength of the peptide-peptide H-bond strength. In the original CG peptide model, (backbone) H-bonds are explicitly modeled using a nonbonded, angular-dependent Lennard-Jones interaction (Bereau and Deserno 2009; Bereau 2011) of the form

$$V_\text{hb}(r,\vartheta_\text{N},\vartheta_\text{C}) = \epsilon_\text{hb}\left[5\left(\frac{\sigma_\text{hb}}{r}\right)^{12} - 6\left(\frac{\sigma_\text{hb}}{r}\right)^{10}\right] \times \begin{cases} \cos^2\vartheta_\text{N}\cos^2\vartheta_\text{C}, & |\vartheta_\text{N}|, |\vartheta_\text{C}| < 90^\circ \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where r is the distance between the two backbone beads N and C′, representing the amide and carbonyl groups, respectively, σ_hb is the equilibrium distance, ϑ_N is the angle formed by the atoms HNC′ (where H is the amide hydrogen) and ϑ_C corresponds to the angle NC′O (where O is the carbonyl oxygen). See illustration in Fig. 1.

The strength of the H-bond potential is modulated by an overall multiplicative prefactor, λ, where $\lambda \leq 1$. The coupling parameter effectively modulates the depth of the interaction potential, impacting the overall stability of H-bonds. HREMD couples copies of the system with different λ values, including the original Hamiltonian, $\lambda = 1$. As will be shown below, reducing the H-bond interaction strength by roughly a factor of 2 is sufficient to lose any trace of helicity.

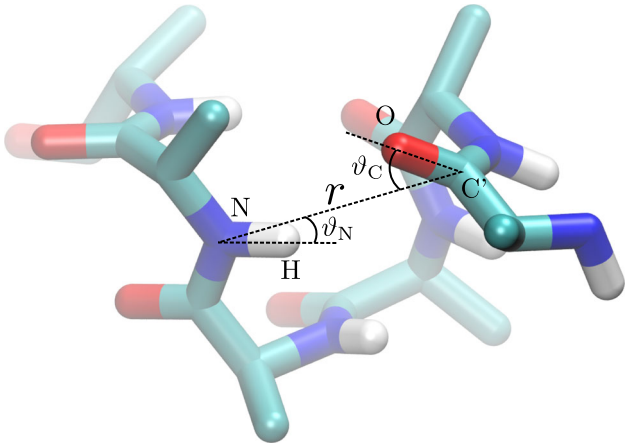

Fig. 1 Geometry of the H-bond interaction between two residues along a peptide chain. Note that the hydrogens were reconstructed geometrically (Bereau and Deserno 2009). Color code: *white* hydrogen; *red* oxygen; *blue* nitrogen; *cyan* carbonyl and alpha carbons, as well as side-chain beads. Simulation snapshot rendered with VMD (Humphrey et al. 1996) (Color figure online)

Table 1 Amino-acid sequences of the peptides studied in this work

Name	Sequence
WALP16	GWWLA LALAL ALAWW A
WALP19	GWWLA LALAL ALALA LWWA
WALP23	GWWLA LALAL ALALA LALAL WWA
fd coat	AEGDD PAKAA FDSLQ ASATE YIGYA WAMVV VIVGA TIGIK LFKKF TSKAS

We thus avoid reducing the coupling parameter to zero, where singularities and numerical instabilities of Lennard-Jones-type potentials often require the use of soft-core alternatives (see, e.g., Gapsys et al. 2012).

Results

We list in Table 1 the amino-acid sequences of all peptides studied in this work.

Folding of WALP16, WALP19, and WALP23

The series of designed WALP peptides has been studied extensively in both simulations (Im and Brooks 2005; Ulmschneider and Ulmschneider 2008; Ulmschneider et al. 2009, 2010; Monticelli et al. 2010; Kim and Im 2010) and experiments (Planque et al. 1998; Killian 2003; Anglin et al. 2009; Holt et al. 2010). WALP consists of alanine-leucine repeat units flanked by tryptophan residues (see Table 1). WALP16 to WALP23 (the number denoting the chain length) serve as excellent model systems to study elementary aspects of transmembrane peptides, such as hydrophobic mismatch via the helix' tilt angle (Planque et al. 1998), but also insertion and folding (Ulmschneider et al. 2010). Focusing on WALP23, we showed previously that our CG model is not biased toward a particular structure: it stabilizes no particular conformation in water, while in the membrane we observed insertion and folding at elevated temperature (i.e., $T = 1.15\,\mathcal{E}/k_B$) (Bereau et al. 2014). Furthermore, WALP's tilt angle distribution, as well as the average helix-helix distance in the membrane, sampled at physiological temperature, agree well with atomistic simulations.

We first compared the structure formation of WALP16 using a plain canonical simulation (i.e., no enhanced sampling) at $T = 1.0\,\mathcal{E}/k_B$ with HREMD at coupling parameters $\lambda \in \{0.5, 0.6, \ldots 1.0\}$. As such, the most altered replica uses an H-bond strength that is halfed compared to the original Hamiltonian, at which point no helicity was sampled according to STRIDE (data not shown). The initial peptide conformation used in both cases is shown in Fig. 2a. Simulations were run for up to 850,000 and 180,000τ for the plain canonical and the HREMD runs, respectively. HREMD swaps between all neighboring pairs of replicas were attempted every 1,000τ. Figure 2c, d show the secondary-structure timeline as a function of residue for both sampling methods. While the lowest replica in the HREMD simulation shows a rapid, persistent increase in helicity, plain canonical sampling leads to a partially helical state (residues 8–16), as illustrated in Fig. 2c. The benefit of HREMD can readily be observed in Fig. 2d, where the progressive increase in helicity contrasts with the long-lived non-native structure sampled from the plain canonical simulation. As a stochastic process, comparing how fast each method performs would require averages over many runs. Though we did not attempt to quantify this difference in the present study, the behavior of the canonical simulation is symptomatic of its propensity to remain trapped in non-native states, while HREMD has allowed us to overcome free-energy barriers and converge toward equilibrium configurations, as further indicated by the examples below. The performance of HREMD over the canonical simulation illustrates the relevance of tuning the H-bond strength itself: long-lived misfolded states are primarily stabilized by non-native H-bonds. By coupling different H-bond-interaction strengths, we help the system hop out of these metastable states.

Similarly, we ran simulations of WALP19, also initialized in a stretched conformation normal to the bilayer, using both plain canonical sampling and HREMD at coupling parameters $\lambda \in \{0.5, 0.6, \ldots 1.0\}$. Figure 3 shows the secondary-structure timelines of both simulations, which run for 1,500,000τ and 120,000τ for the plain canonical and HREMD runs, respectively. We also find a quick,

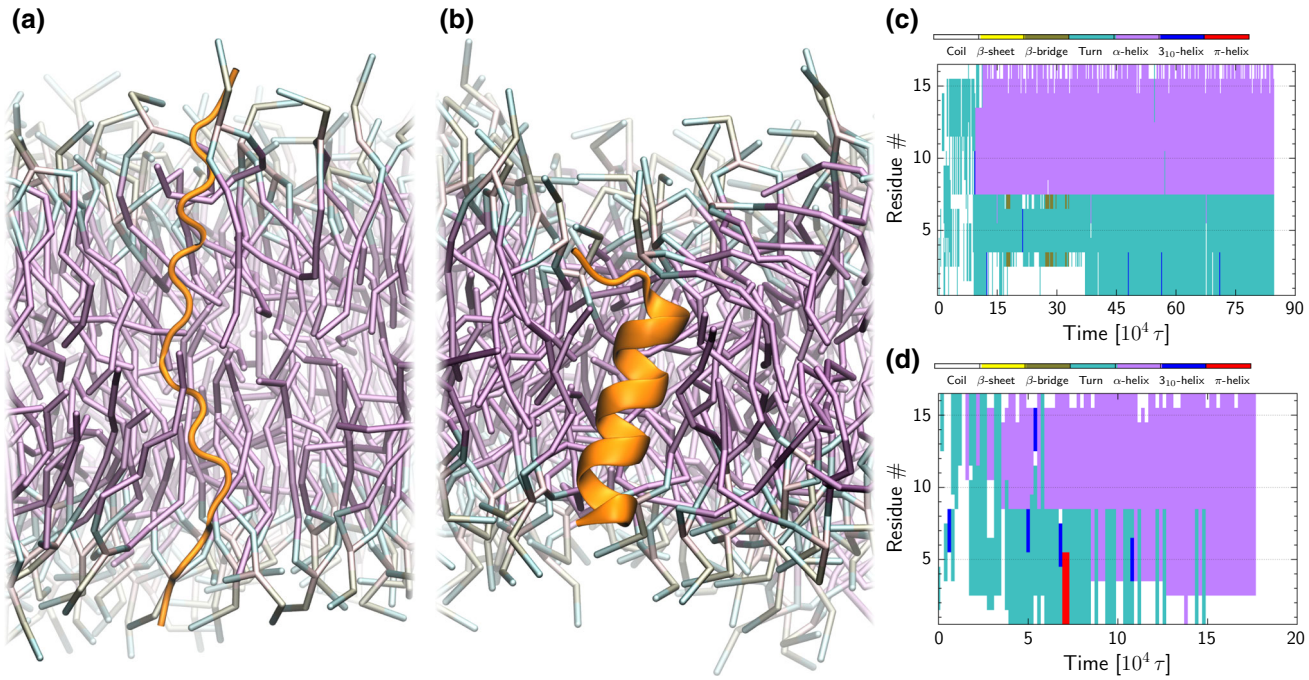

Fig. 2 Folding of WALP16 in a POPC membrane. Cartoon representations of **a** initial, unstructured ($t = 0$) and **b** equilibrium, folded conformation (sampled from HREMD; $t = 150{,}000\tau$). The peptide is depicted in *orange*, where *thick and thin ribbons* correspond to the helical and coil states, respectively; the lipids are color-coded according to their bead type: *purple* for the hydrocarbon chains, *light pastel colors* for the interfacial and head groups. Timeline of secondary structure as a function of residue for **c** the canonical simulation (i.e., no enhanced sampling) and **d** HREMD—note the difference in simulation timescales. *White, green, and purple* events in the timeline correspond to coil, turn, and α-helical conformations, respectively (Color figure online)

initial, helix formation for both protocols, though the simulation without enhanced sampling spends significant time sampling a partial 7-residue-long helix, as the rest of the peptide only samples turn and coil structures (Fig. 3a), until $1{,}200{,}000\tau$. The homogeneity of the secondary-structure timeline during folding highlights how the system can get stuck in a kinetic trap for long times. On the other hand, HREMD shows an alternation between structures forming an α-helix at different positions along the chain. Excluding fluctuations at the termini, the helix finally extends throughout the chain to form the folded transmembrane conformation at around $t = 70{,}000\tau$ (Fig. 3b).

Turning to WALP23, we only investigated folding using HREMD. A simulation of $150{,}000\tau$ using the same coupling parameters as for WALP19 indeed reached the transmembrane, fully helical state at 310 K (data not shown). This not only highlights the robustness of the method to equilibrate simple helical peptides in the membrane, but it also confirms that the helical transmembrane state of WALP23 is indeed the equilibrium state at physiological temperature for this CG model—a claim we previously (Bereau et al. 2014) supported by indirect evidence but could not demonstrate directly.

To understand better how the peptide affects the shape of the membrane during the folding process, we monitored the distance from the bilayer midplane to each glycerol bead—a proxy for the monolayer thickness, h, as a function of its lateral distance to the peptide's center of mass, as illustrated in Fig. 4a. Since we are specifically interested in how *partially* folded structures affect the membrane, both initially stretched, unstructured (i.e., no helicity), and late-stage folded (i.e., helicity ratio larger than 75 %) conformations were discarded, and the data were averaged over the remaining snapshots. The $h(d)$ profiles of each WALP peptide are shown in Fig. 4b. They illustrate that the free energy of partially folded peptides must also contain an elastic contribution derived from local bending of the host membrane's monolayers. We observe that the peptide-induced membrane deformation relaxes over $d \approx 20$ Å. This length is comparable to monolayer thickness and suggests that the associated energetics is best described by local lipid physics (packing, stretching, tilting, etc.; see, e.g., (Lundbæk et al. 2010) and references therein). A similar relaxation length scale and profile was observed from CG simulations of cylindrical model peptides in the context of hydrophobic mismatch (Venturoli et al. 2005). Remarkably, we also observe a small dip at significantly longer length scales (i.e., $d \simeq 60$ Å), suggestive of curvature elasticity effects, with their often characteristic damped oscillatory relaxation. The use of a significantly

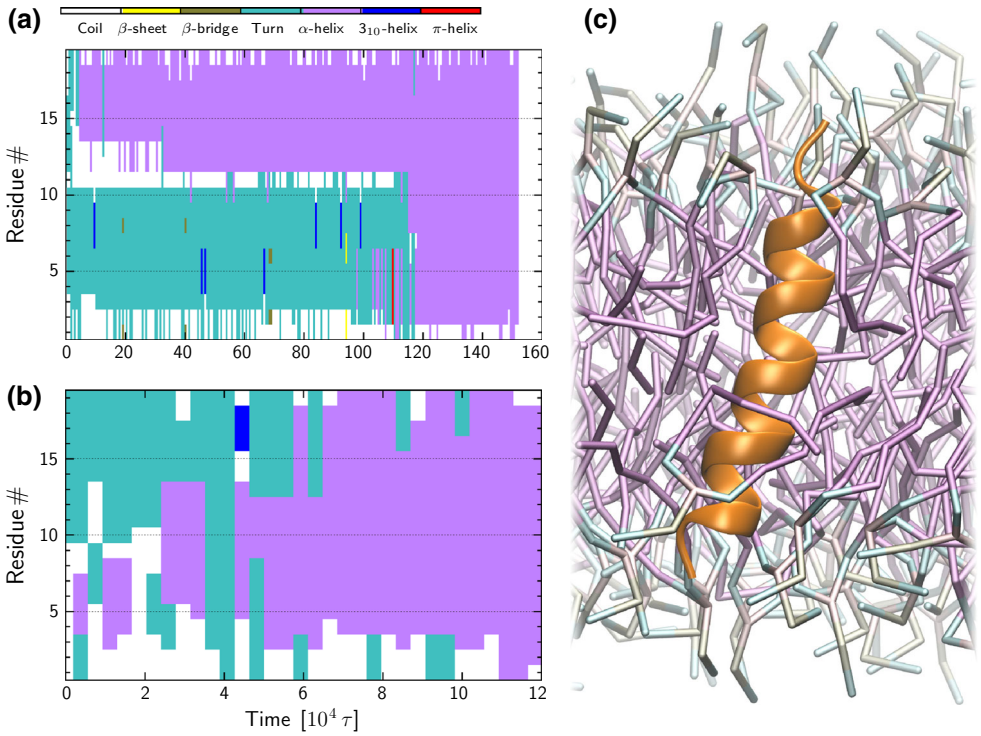

Fig. 3 Folding of WALP19 in a POPC membrane. Timeline of secondary structure as a function of residue for **a** the canonical simulation (i.e., no enhanced sampling) and **b** HREMD—note the difference in simulation timescales. At the end of the runs, the plain canonical and the HREMD simulations folded to the equilibrium state. *White, green, and purple* events in the timeline correspond to coil, turn, and α-helical conformations, respectively. **c** Final conformation from the HREMD simulation (Color figure online)

smaller box size would enforce (i) no long-range oscillation behavior and (ii) a shorter decay, leading to larger deformation energies. We find here a correlation between the length of the peptide and the deformation profile (i.e., the shorter the peptide the larger the deformation), though statistics over several independent runs would be necessary to confirm this behavior. Additional simulations (data not shown) indicate that although the overall shapes shown in Fig. 4b are representative of the membrane's behavior, the offset deformation close to the peptide depends sensitively on the folding process, and thus would also require averaging over many folding experiments. Irrespective of a quantitative analysis, a pronounced membrane thinning close to the peptide is to be expected due to the role played by hydrophobic mismatch (Fattal and Ben-Shaul 1993): The WALP peptides studied here contain hydrophobic segments that are comparable in length to the membrane thickness. The formation of any additional turn that ultimately shapes the α-helix will require significant backbone distortion, thereby further shortening the distance between the two polar endgroups, compared to the hydrophobic region of the membrane. We would like to remind the reader that the well known mattress model of Mouritsen and Bloom (Mouritsen and Bloom 1984) describes the membrane-mediated interaction of transmembrane helices that induce hydrophobic mismatch in a way that is mathematically similar to springs spanning through (and locally pinching) a mattress. It is worthwhile

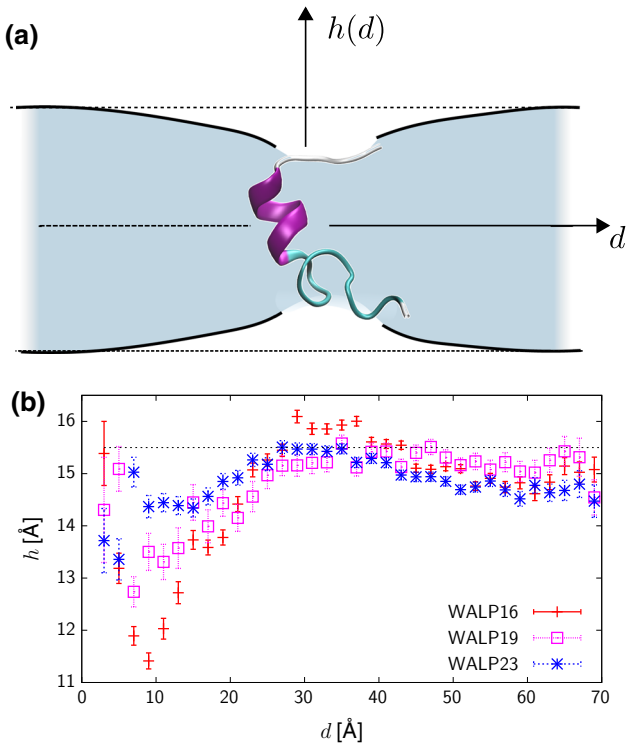

Fig. 4 **a** Schematic representation of the glycerol bead height h above the bilayer midplane as a function of its lateral distance, d, measured from the peptide's center of mass. **b** $h(d)$ for the three WALP peptides. *Error bars* denote the error of the mean. The *horizontal, dashed line* is a mere guide to the eye representing the unperturbed POPC bilayer

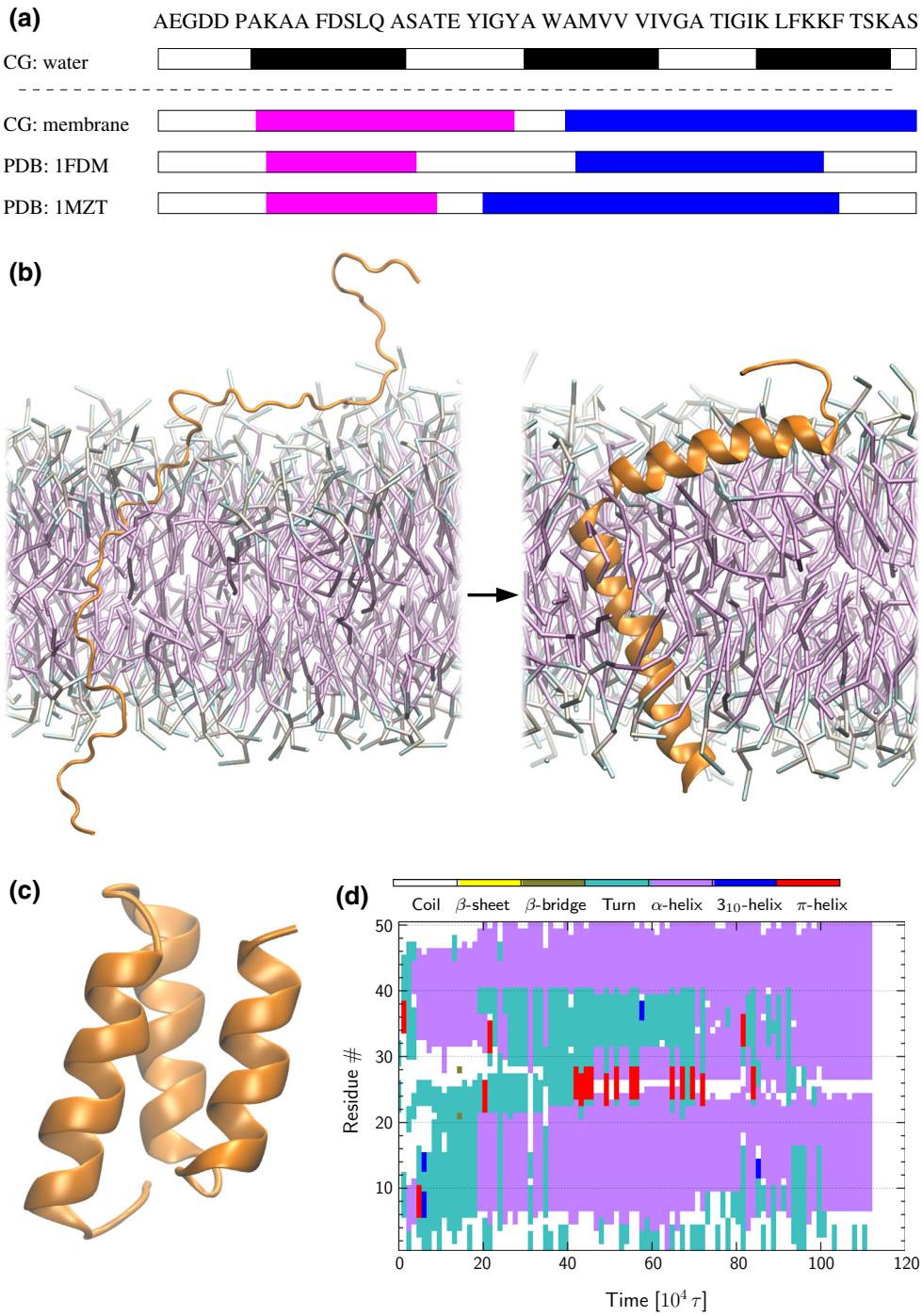

Fig. 5 Folding of the fd coat protein. **a** Helical residues from the CG simulation in water, in the membrane, and PDBs 1FDM (Almeida and Opella 1997) and 1MZT (Marassi and Opella 2003). *Any filled bar* represents a helical region (helical state during at least 80 % of the simulation): *blue* for transmembrane, *magenta* for interfacial, and *black* otherwise. **b** Early, unstructured and equilibrium ($t = 1{,}100{,}000\tau$) conformations of the CG simulation with the membrane. **c** Equilibrium structure of the CG simulation in water (the bead representation is hidden for clarity). **d** Timeline of secondary structure as a function of residue in the membrane. *White, green, and purple* events in the timeline correspond to coil, turn, and α-helical conformations, respectively (Color figure online)

to remember that in our case these springs are loaded during the alpha-helical folding process, and that, therefore, this loading energy must derive from—and is limited by—the free energy available from secondary structure formation. Interestingly, we found qualitatively similar profiles when focusing on membrane deformations with the folded peptides (i.e., helicity ratio larger than 75 %), though the features were toned down (data not shown).

Folding of the fd Coat Protein

The last peptide we study is the major pVIII coat protein (fd coat) of the filamentous fd bacteriophage. This 50-residue-long protein, which can be found within the membrane of infected bacteria prior to virus assembly, has been investigated experimentally by both solution NMR in detergent micelles (PDB: 1FDM) (Almeida and Opella 1997) and solid-state NMR (PDB: 1MZT) (Marassi and

Opella 2003). Im and Brooks later folded atomistically fd coat in implicit water and membrane environments (Im and Brooks 2004). In all cases, two distinct helices were identified, one interfacial and one transmembrane, joined by a short loop in the vicinity of residue 20.

We ran HREMD simulations at coupling parameters $\lambda \in \{0.75, 0.80, \ldots 1.00\}$ for 1,100,000τ, starting from a stretched, unstructured configuration spanning the bilayer normal. To control against any potential force-field bias from the CG model, we also ran simulations of the same protein in the (implicit) aqueous environment (i.e., without the membrane) using a standard parallel tempering algorithm (Bereau and Deserno 2009) with six temperatures between $T = 1.00\,\mathcal{E}$ and $1.30\,\mathcal{E}$ for t =200,000τ. Averaging over the last 50,000τ, we extracted the per-residue helical ratio of the protein in each environment. Any residue identified as helical during at least 80 % of the simulation (at $T = 1.00\,\mathcal{E}/k_B$) was marked as such—the results for both simulations are shown in Fig. 5a. The two simulations stabilize significantly different folds: In water, we observe three helices connected by loops (fraction of helicity for each loop-residue is less than 10 %), formed around the glycine residues at positions 23, as well as 34 and 38. The helices tightly pack to minimize hydrophobic exposure to the surface, see simulation snapshot in Fig. 5c. On the other hand, the CG simulation in the membrane shows only two helices (Fig. 5a, b, d), while the other regions showed negligible helical content (as shown in the timeline of Fig. 5d). We find that the loop between the two helices in the membrane simulation occurs at residues 25–28, while those are helical in the water simulation. As such, we find that the CG model stabilizes different structures in the water and membrane environments, as we had observed previously for WALP23 (Bereau et al. 2014). We refrain from drawing any further conclusion from the water simulation, as its accuracy is unclear: atomistic simulations have not been performed to the best of our knowledge, and any experimental result would include finite-concentration effects, likely driving aggregation.

The peptide structure in the membrane consists of four separate regions: an unstructured N-terminal tail dangling above the membrane (Ala1–Pro6); the interfacial helix (Ala7–Tyr24); a loop connecting the two helices (Ala25–Trp26); and the transmembrane helix (Ala27–Lys48). As seen in Fig. 5a, this compares favorably with both 1FDM and 1MZT experimental structures. Moreover, Im and Brooks' atomistic simulations also folded two helices that shared common traits with both experimental structures in their length and relative orientation. While the exact location of the turn region between the two helices may require further refinement, we observe that the CG model can fold and stabilize the two secondary-structure motifs at roughly the right location in the sequence. Moreover, it illustrates the robustness of the CG parametrization: while weaker hydrogen bonds would lead to little or no helix formation, stronger hydrogen bonds would likely give rise to one long transmembrane α-helix, or at least an over-stabilization of the helical extent. Note that our loop is shorter than 1FDM, but longer than 1MZT. The presence of the turn to accommodate for both the transmembrane and the interfacial helices shows that the partitioning of amino acids between water and the bilayer are, at least to a certain extent, well imprinted in the CG model, and well balanced with its hydrogen-bond strength.

Conclusion

Structure formation in the membrane constitutes a major challenge of sampling in computer simulations. By merely tuning the strength of backbone H-bonds, the HREMD protocol aims at shortening the lifetime of long-lived non-native metastable states. The few degrees of freedom altered within the HREMD scheme limit the number of replicas required. Moreover, the weak stability of H-bonds implies that a relatively small decrease in the H-bond strength prevents the formation and stabilization of secondary structure, such that the coupling parameter λ need not be brought all the way to zero. We illustrated the improved equilibration capabilities of this HREMD protocol by comparing it with simulations without enhanced sampling for the structure formation of WALP16 and WALP19. The folding of WALP19 without HREMD demonstrated the efficiency of the enhanced-sampling protocol and illustrated how a membrane-bound peptide spends significant times in kinetic traps. This methodology also allowed us to equilibrate WALP23 from an unstructured conformation to its transmembrane helical state. In addition to enhanced peptide-sampling capabilities, we emphasize the absence of structural restraints on the lipids, as opposed to previous atomistic simulations (Nymeyer et al. 2005). Likewise, we do not rely on higher-temperature simulations to reach equilibrium (Ulmschneider et al. 2010; Bereau et al. 2014). We highlight the agreement with both experimental results and atomistic simulations in the structure of the fd coat protein, which consisted of both interfacial and transmembrane helices. The folding of this long protein (50 residues) required significant sampling at the CG level, and was only tackled atomistically using an implicit membrane model. The strengths of explicitly modeling the membrane—compared to a dielectric-slab approach—were illustrated by monitoring peptide-induced membrane-thickness deformation, as well as the previously reported transient pore formation due to the cooperative effect of interfacial peptides (Bereau et al. 2014). Beyond the robustness of the sampling method, this study illustrates

the balance in the CG parametrization between hydrogen-bond strength and amino-acid partioning between water and the bilayer. It also serves as a demonstration of the ability for CG models to tackle structure formation in the membrane. While such an approach is not meant to replace atomistic simulations of peptides in the membrane, it can much more efficiently sample phase space (we remind the reader that the reference atomistic simulation used an implicit membrane model) and identify relevant conformations, which one can later backmap for subsequent atomistic simulations (Bereau et al. 2012).

Acknowledgments We thank Zun-Jing Wang for her careful work parametrizing and implementing the lipid CG model, as well as early validation of the peptide–lipid cross-parametrization. We are also indebted to Cameron Abrams, Mike Allen, Frank Brown, Ed Lyman, and Alex Sodt, for constructive suggestions on our methodologies. We acknowledge Aoife Fogarty, Kurt Kremer, and Debashish Mukherji for critical reading of the manuscript. This work was partially funded by the Max Planck Institute for Polymer Research (TB) and NSF Grant MCB #1330226 (MD).

References

Affentranger R, Tavernelli I, Di Iorio EE (2006) A novel hamiltonian replica exchange md protocol to enhance protein conformational space sampling. J Chem Theory Comput 2(2):217–228

Almeida F, Opella S (1997) fd coat protein structure in membrane environments: structural dynamics of the loop between the hydrophobic trans-membrane helix and the amphipathic in-plane helix. J Mol Biol 270(3):481–495

Anglin TC, Brown KL, Conboy JC (2009) Phospholipid flip-flop modulated by transmembrane peptides WALP and Melittin. J Struct Biol 168(1):37–52

Baaden M, Marrink SJ (2013) Coarse-grain modelling of protein–protein interactions. Curr Opin Struct Biol 23(6):878–886

Bereau T (2011) Unconstrained structure formation in coarse-grained protein simulations. Ph.D. thesis, Carnegie Mellon University

Bereau T, Bachmann M, Deserno M (2010) Interplay between secondary and tertiary structure formation in protein folding cooperativity. J Am Chem Soc 132(38):13,129–13,131

Bereau T, Deserno M (2009) Generic coarse-grained model for protein folding and aggregation. J Chem Phys 130(23):235,106

Bereau T, Deserno M, Bachmann M (2011) Structural basis of folding cooperativity in model proteins: insights from a microcanonical perspective. Biophys J 100(11):2764–2772

Bereau T, Globisch C, Deserno M, Peter C (2012) Coarse-grained and atomistic simulations of the salt-stable cowpea chlorotic mottle virus (SS-CCMV) subunit 26–49: β-barrel stability of the hexamer and pentamer geometries. J Chem Theory Comput 8(10):3750–3758

Bereau T, Wang ZJ, Deserno M (2014) More than the sum of its parts: coarse-grained peptide–lipid interactions from a simple cross-parametrization. J Chem Phys 140(11):115,101

Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) The protein data bank. Eur J Biochem 80(2):319–324

Bunker A, Dünweg B (2000) Parallel excluded volume tempering for polymer melts. Phys Rev E 63(1):016,701

Caffrey M (2003) Membrane protein crystallization. J Struct Biol 142(1):108–132

DeLucas L (2009) Membrane protein crystallization, vol 63. Academic Press, San Diego

de Planque MR, Greathouse DV, Koeppe RE, Schäfer H, Marsh D, Killian JA (1998) Influence of lipid/peptide hydrophobic mismatch on the thickness of diacylphosphatidylcholine bilayers. a 2H NMR and ESR study using designed transmembrane α-helical peptides and gramicidin A. Biochemistry 37(26):9333–9345

Fattal DR, Ben-Shaul A (1993) A molecular model for lipid–protein interaction in membranes: the role of hydrophobic mismatch. Biophys J 65(5):1795

Filipe HA, Moreno MJ, Róg T, Vattulainen I, Loura LM (2014) How to tackle the issues in free energy simulations of long amphiphiles interacting with lipid membranes: convergence and local membrane deformations. J Phys Chem B 118(13):3572–3581

Freddolino PL, Harrison CB, Liu Y, Schulten K (2010) Challenges in protein-folding simulations. Nature Phys 6(10):751–758

Freddolino PL, Park S, Roux B, Schulten K (2009) Force field bias in protein folding simulations. Biophys J 96(9):3772–3780

Frishman D, Argos P (1995) Knowledge-based protein secondary structure assignment. Proteins: Struct Funct Bioinf 23(4):566–579

Gapsys V, Seeliger D, de Groot BL (2012) New soft-core potential function for molecular dynamics based alchemical free energy calculations. J Chem Theory Comput 8(7):2373–2382

Globisch C, Krishnamani V, Deserno M, Peter C (2013) Optimization of an elastic network augmented coarse grained model to study CCMV capsid deformation. PLoS One 8(4):e60,582

Grossfield A, Feller SE, Pitman MC (2007) Convergence of molecular dynamics simulations of membrane proteins. Proteins: Struct Funct Bioinf 67(1):31–40

Holt A, Rougier L, Réat V, Jolibois F, Saurel O, Czaplicki J, Killian JA, Milon A (2010) Order parameters of a transmembrane helix in a fluid bilayer: case study of a walp peptide. Biophys J 98(9):1864–1872

Huang J, MacKerell AD Jr (2014) Induction of peptide bond dipoles drives cooperative helix formation in the $(AAQAA)_3$ peptide. Biophys J 107(4):991–997

Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14(1):33–38

Im W, Brooks CL (2005) Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations. Proc Natl Acad Sci USA 102(19):6771–6776

Im W, Brooks CL III (2004) De novo folding of membrane proteins: an exploration of the structure and NMR properties of the fd coat protein. J Mol Biol 337(3):513–519

Ingólfsson HI, Lopez CA, Uusitalo JJ, de Jong DH, Gopal SM, Periole X, Marrink SJ (2014) The power of coarse graining in biomolecular simulations. Wiley Interdiscip Rev 4(3):225–248

Jiang W, Roux B (2010) Free energy perturbation hamiltonian replica-exchange molecular dynamics (FEP/H-REMD) for absolute ligand binding free energy calculations. J Chem Theory Comput 6(9):2559–2565

Kar P, Feig M (2014) Recent advances in transferable coarse-grained modeling of proteins. Adv Protein Chem Struct Biol

Kar P, Gopal SM, Cheng YM, Panahi A, Feig M (2014) Transferring the primo coarse-grained force field to the membrane environment: simulations of membrane proteins and helix–helix association. J Chem Theory Comput 10(8):3459–3472

Killian JA (2003) Synthetic peptides as models for intrinsic membrane proteins. FEBS Lett 555(1):134–138

Kim T, Im W (2010) Revisiting hydrophobic mismatch with free energy simulation studies of transmembrane helix tilt and rotation. Biophys J 99(1):175–183

Kolb A, Dünweg B (1999) Optimized constant pressure stochastic dynamics. J Chem Phys 111(10):4453–4459

Limbach HJ, Arnold A, Mann BA, Holm C (2006) Espresso: an extensible simulation package for research on soft matter systems. Comput Phys Commun 174(9):704–727

Liu P, Huang X, Zhou R, Berne B (2006) Hydrophobic aided replica exchange: an efficient algorithm for protein folding in explicit solvent. J Phys Chem B 110(38):19,018–19,022

Liu P, Kim B, Friesner RA, Berne B (2005) Replica exchange with solute tempering: a method for sampling biological systems in explicit water. Proc Natl Acad Sci USA 102(39):13,749–13,754

Lundbæk JA, Collingwood SA, Ingólfsson HI, Kapoor R, Andersen OS (2010) Lipid bilayer regulation of membrane protein function: gramicidin channels as molecular force probes. J R Soc Interface 7(44):373–395

Lyman E, Ytreberg FM, Zuckerman DM (2006) Resolution exchange simulation. Phys Rev Lett 96(2):028,105

MacCallum JL, Bennett W, Tieleman DP (2008) Distribution of amino acids in a lipid bilayer from computer simulations. Biophys J 94(9):3393–3404

Marassi FM, Opella SJ (2003) Simultaneous assignment and structure determination of a membrane protein from NMR orientational restraints. Protein Sci 12(3):403–411

Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink SJ (2008) The martini coarse-grained force field: extension to proteins. J Chem Theory Comput 4(5):819–834

Monticelli L, Tieleman DP, Fuchs PF (2010) Interpretation of ^{2}H-NMR experiments on the orientation of the transmembrane helix walp23 by computer simulations. Biophys J 99(5):1455–1464

Mori T, Jung J, Sugita Y (2013) Surface-tension replica-exchange molecular dynamics method for enhanced sampling of biological membrane systems. J Chem Theory Comput 9(12):5629–5640

Mouritsen OG, Bloom M (1984) Mattress model of lipid–protein interactions in membranes. Biophys J 46(2):141–153

Mu Y (2009) Dissociation aided and side chain sampling enhanced hamiltonian replica exchange. J Chem Phys 130(16):164,107

Neale C, Bennett WD, Tieleman DP, Pomès R (2011) Statistical convergence of equilibrium properties in simulations of molecular solutes embedded in lipid bilayers. J Chem Theory Comput 7(12):4175–4188

Neale C, Madill C, Rauscher S, Pomès R (2013) Accelerating convergence in molecular dynamics simulations of solutes in lipid membranes by conducting a random walk along the bilayer normal. J Chem Theory Comput 9(8):3686–3703

Noid W (2013) Perspective: coarse-grained models for biomolecular systems. J Chem Phys 139(9):090,901

Nymeyer H, Woolf TB, Garcia AE (2005) Folding is not required for bilayer insertion: replica exchange simulations of an α-helical peptide with an explicit lipid bilayer. Proteins Struct Funct Bioinf 59(4):783–790

Okabe T, Kawata M, Okamoto Y, Mikami M (2001) Replica-exchange Monte Carlo method for the isobaric–isothermal ensemble. Chem Phys Lett 335(5):435–439

Paloncýová M, Berka K, Otyepka M (2012) Convergence of free energy profile of coumarin in lipid bilayer. J Chem Theory Comput 8(4):1200–1211

Paschek D, García AE (2004) Reversible temperature and pressure denaturation of a protein fragment: a replica exchange molecular dynamics simulation study. Phys Rev Lett 93(23):238,105

Periole X, Cavalli M, Marrink SJ, Ceruso M (2009) Combining an elastic network with a coarse-grained molecular force field: structure, dynamics and intermolecular recognition. J Chem Theory Comput 5:2531–2543

Piana S, Lindorff-Larsen K, Shaw DE (2011) How robust are protein folding simulations with respect to force field parameterization? Biophys J 100(9):L47–L49

Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 314(1):141–151

Swendsen RH, Wang JS (1986) Replica Monte Carlo simulation of spin glasses. Phys Rev Lett 57(21):2607–2609

Ulmschneider JP, Doux JP, Killian JA, Smith JC, Ulmschneider MB (2009) Peptide partitioning and folding into lipid bilayers. J Chem Theory Comput 5(9):2202–2205

Ulmschneider MB, Doux JP, Killian JA, Smith JC, Ulmschneider JP (2010) Mechanism and kinetics of peptide partitioning into membranes from all-atom simulations of thermostable peptides. J Am Chem Soc 132(10):3452–3460

Ulmschneider MB, Ulmschneider JP (2008) Folding peptides into lipid bilayer membranes. J Chem Theory Comput 4(11): 1807–1809

Venturoli M, Smit B, Sperotto MM (2005) Simulation studies of protein-induced bilayer deformations, and lipid-induced protein tilting, on a mesoscopic model for lipid bilayers with embedded proteins. Biophys J 88(3):1778–1798

von Heijne G (2011) Introduction to theme membrane protein folding and insertion. Annu Rev Biochem 80:157–160

Wan CK, Han W, Wu YD (2011) Parameterization of pace force field for membrane environment and simulation of helical peptides and helix–helix association. J Chem Theory Comput 8(1): 300–313

Wang ZJ, Deserno M (2010) Systematic implicit solvent coarse-graining of bilayer membranes: lipid and phase transferability of the force field. New J Phys 12(9):095,004

Wang ZJ, Deserno M (2010) A systematically coarse-grained solvent-free model for quantitative phospholipid bilayer simulations. J Phys Chem B 114(34):11,207–11,220