

# Hybrid Particle-Field Model for Conformational Dynamics of Peptide Chains

Sigbjørn Løland Bore,<sup>†</sup> Giuseppe Milano,<sup>‡</sup> and Michele Cascella<sup>\*,†</sup>

<sup>†</sup>*Department of Chemistry, and Hylleraas Centre for Quantum Molecular Sciences,  
University of Oslo, PO Box 1033 Blindern, 0315 Oslo, Norway*

<sup>‡</sup>*Department of Organic Materials Science, University of Yamagata, 4-3-16 Jonan  
Yonezawa, Yamagata-ken 992-8510, Japan*

E-mail: michele.cascella@kjemi.uio.no

## Abstract

We propose the first model of a polypeptide chain based on a hybrid-particle field approach. The intramolecular potential is built on a two-bead coarse grain mapping for each amino acid. We employ a combined potential for the bending and the torsional degrees of freedom that ensures the stabilization of secondary structure elements in the conformational space of the polypeptide. The electrostatic dipoles associated to the peptide bonds of the main-chain are reconstructed by a topological procedure. The intermolecular interactions comprising both the solute and the explicit solvent are treated by a density functional-based mean-field potential. Molecular dynamics simulations on a series of test systems show how the model here introduced is able to capture all the main features of polypeptides. In particular, homopolymers of different lengths yield a complex folding phase diagram, covering from the collapsed to swollen state. Moreover, simulations on models of a four-helix bundle and of an alpha+beta peptide evidence how the collapse of the hydrophobic core drives the appearance of both folded motifs and the stabilization of tertiary or quaternary assemblies. Finally,

the polypeptide model is able to structurally respond to the environmental changes caused by the presence of a lipid bilayer.

## 1 Introduction

Field-based approaches where molecules are represented by density-fields offer today a rigorous and computationally efficient way of simulating condensed-phase systems at the mesoscopic scale.<sup>1</sup> Particularly popular in polymer science are the self consistent field (SCF) approaches originating from the work by Edwards<sup>2</sup> and Helfand.<sup>3,4</sup> In SCF, the polymer segments composing a melt are treated as non interacting systems coupled to an external field responsible for the deviation from the ideal behavior. In analogy to Kohn-Sham equations for quantum density functional theory,<sup>5</sup> the external field is computed self consistently from a statistical average of the spatially inhomogeneous segment density distributions interacting with the external field. Applications ranging from polymer brushes,<sup>6,7</sup> polymer composites,<sup>8</sup> block copolymers melts,<sup>9,10</sup> to colloid particles,<sup>11</sup> demonstrate the usefulness of the method.

The central step in density-based approaches is the definition of a mesoscopic density associated to a microscopic molecular configurations of the system. For molecules of complex architecture, this step can be challenging. Recently, Müller et al<sup>12</sup> proposed a hybrid particle-field scheme (hPF hereafter) which retained computational advantages of SCF-theory while keeping the molecular detail of the polymer chains. In this approach, a polymeric system is modeled as an ensemble of independent chains in a fluctuating external field. The external field models instantaneous interactions between chains and is computed from spatial inhomogeneous densities of particles. Configurations of the system are sampled through a Monte Carlo procedure.<sup>13</sup> hPF methods has been successfully applied on various polymer and soft matter systems including homopolymers, block-polymers or nanocomposites<sup>13-18</sup>.

The extension of the hPF formalism to the calculation of density gradients and forces led to the possibility of coupling the hPF method with molecular dynamics

protocols (hPF-MD).<sup>19</sup> Calculation of potential energy terms and forces from local densities results in much more efficient simulation than ordinary molecular dynamics.<sup>20</sup> Firstly, by computing forces from the density field, the expensive non-bonded pair interactions of ordinary MD are avoided. Secondly, by only storing the density and not pair list in the memory, the time propagation algorithm for the motion of the individual particles is *de facto* embarrassingly parallel.<sup>20</sup>

First application studies on homopolymers and block copolymers<sup>19,21</sup> demonstrated the effectiveness of hPF-MD in describing soft-matter and polymeric materials both at the mesoscale and atomistic resolution.<sup>22</sup> More recently, hPF-MD was used to describe the morphology and dynamics of amphiphilic systems in the aqueous environment, in particular of phospholipids<sup>23-26</sup> and surfactants,<sup>27,28</sup> paving the way to the possibility of using hPF-MD to investigate biological systems.<sup>29</sup>

Among the biological polymers, the development of computational multi-scaling approaches for proteins is particularly challenging, but of great interest. In fact, biologically active polypeptide chains can be as short as few tens of amino acids (aa hereafter), or comprise single chains of tens of thousands of aa, like for example in titin.<sup>30</sup> Moreover, individual proteins can aggregate or assemble in organized multi-mers yielding nanofibers and nanotubules,<sup>31,32</sup> as well as give rise to crowded functional structures, for example, in several membrane-interfering antimicrobial peptides.<sup>33</sup> The computational investigation of such phenomena at a computational level relies on the ability of models to describe both the local interactions determining structural features of each protein-protein assembly, and the interplay among different proteins and the environment that yields to the formation of complex nanostructures.

Molecular-resolution approaches based on parameterized potentials, and in particular the representation at all-atom resolution, are in principle the best models for describing molecular interactions.<sup>34</sup> Unfortunately, their computational costs are still too high to allow their use for the investigation of the dynamics of protein complexes at the mesoscale.

Here, we introduce a first hPF-MD approach to study protein dynamics. hPF-

MD requires an underlying molecular model at coarse-grained (CG) resolution. To ensure good performance and high accuracy of hPF-MD, the CG model has to possess a series of features. First, the intra-molecular potential driving the conformational changes of the molecule has to reproduce, at the pure CG level, the main features of interest of the conformational space. Second, the CG mapping should represent all the molecular moieties with a similar resolution, with the fewest possible bead types. This is necessary in order to have numerical stability with the grid size used for the density evaluation, and the highest computational efficiency for hPF-MD.

The simplest approach to model a protein can be one single ellipsoid or spherical bead which is matched to the hydrodynamic radius of the overall chain. This approach has been applied to study phenomena on very large diffusion scales where the protein can be approximated by a single diffusing entity.<sup>35</sup> Retaining a molecular-detailed description, several CG potentials have been proposed in the past,<sup>36-39</sup> aiming for a direct physical representation of the interactions driving the motion of the peptide chains. The functional forms and the corresponding parameters available today are derived from either bottom-up or top-down strategies,<sup>40</sup> by calibrating parameters on thermodynamic experimental data like partition coefficients in water/oil phases,<sup>41,42</sup> or from higher-resolution all-atom simulations.<sup>43,44</sup> For example, force-matching techniques have been developed by Voth and coworkers<sup>45-47</sup> and successfully been applied on peptides,<sup>48</sup> protein assemblies<sup>49,50</sup> and protein-membrane interactions.<sup>51,52</sup>

The definition of the intra-molecular terms is a key issue for the reliability of CG models in describing the conformational landscape of a polypeptide. In fact, some of the most successful potentials either make use of molecular Hamiltonians characterized by very complex functional forms<sup>53</sup> or adopt simpler potential energy functions where the local configurations and secondary structure elements are restrained to a given reference.<sup>41</sup> At present, it is understood that, in order to correctly describe the complexity of the conformational landscape of polypeptides, the introduction of directional contact interactions mimicking the formation of H-bonds between peptide bonds is required.<sup>37,54,55</sup>

Explicit treatment of electrostatic interactions is usually avoided in CG simulations,<sup>37</sup> due to its very high computational costs in the condensed phase.<sup>56</sup> In fact, long-range electrostatics is a key ingredient of the physics of proteins.<sup>57</sup> Several CG models for proteins have progressively introduced it at different levels of complexity.<sup>41,58–66</sup>

Among others, in the past decade Cascella and co-workers introduced a model where the orientation of the backbone dipoles are directly reconstructed starting from the geometry of the  $C\alpha$ 's.<sup>67</sup> This protocol gives CG electrostatics in agreement with atomistic simulations,<sup>67</sup> and is able to stabilize protein secondary structures.<sup>68,69</sup> Starting from this approach, here we propose a poly-peptide model based on a two-bead mapping, where the hPF-MD formalism is used to describe intermolecular interactions. The functional form of the intra-molecular terms for this model is designed to describe the complex conformational landscape of peptide chains.

Using a series of test cases, we show how the presented model represents well the main physics of a polymer chain, including the existence of marginally stable folded regions between collapsed and swollen structures, the stabilization of folded structures by formation of a hydrophobic core, and the occurrence of local and global structural changes upon variation of the environmental conditions.

## 2 The Polypeptide Model and its Potential Function

### 2.1 Coarse-Grained structural model of the polypeptide chain

We introduce here a *two bead* CG model (Figure 1), where each aa composing the peptide chain is represented by one bead centered at its  $C_\alpha$  (CA, hereafter), and one bead for its side chain (CB). Additionally, one electrostatic dipole  $\mu$  that mimics the presence

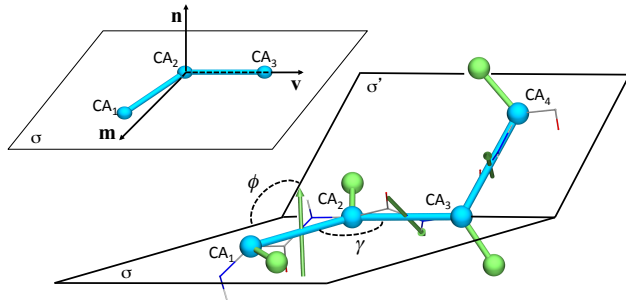


Figure 1: Model of the polypeptide chain. *Top*: Definition of the internal coordinate system of three consecutive CAs used for the reconstruction of the CB beads and the peptide dipoles. *Bottom*: Constitutional elements of the CG model for the peptide chain. the CA beads mapping the all-atom polypeptide main-chain (thin licorice) are in cyan spheres. The CB beads representing the side-chains of the amino acids are depicted in green spheres. The green arrows represent the electrostatic dipoles mimicking the peptide bond moieties. Both CB positions and dipole orientations are determined from the positions of the CAs. The figure shows the bending angle  $\gamma$  and the torsional angle  $\phi$  entering as variables in the combined bending-torsional potential.

of the peptide bond is added at the midpoint of each CA-CA bond. The positions of the CA beads are the only degrees of freedom of the systems. This mapping has some analogy to the UNRES model,<sup>53</sup> where an effective term is used to map the Hydrogen-bond interactions at contact distance between peptide moieties. The location and orientations of the CB's and the dipoles are determined from the coordinates of the CA's, according to the procedure described in the following paragraphs.

**Internal coordinate system** Following ref.,<sup>67</sup> given three consecutive CA's in a peptide chain, we define a local reference system with orthonormal basis vectors  $\hat{\mathbf{v}}$ ,  $\hat{\mathbf{n}}$  and  $\hat{\mathbf{m}}$  (shown in Figure 1) with its origin at the central CA<sub>2</sub> bead. The first unit vector  $\hat{\mathbf{v}}$  is oriented along the CA<sub>2</sub>-CA<sub>3</sub> bond, the second unit vector  $\hat{\mathbf{n}}$  is oriented along the normal direction to the CA<sub>1</sub>-CA<sub>2</sub>-CA<sub>3</sub> plane, and the third orthogonal unit vector  $\hat{\mathbf{m}}$  is determined from the first two using the right-hand rule.

**Dipole reconstruction** The dipoles of the backbone  $\{\boldsymbol{\mu}\}$  are represented by two charges  $q_{\pm} = \pm 0.25$  au positioned at:

$$\mathbf{r}_{\pm}(\gamma) = \frac{1}{2}\mathbf{r}_3 \pm \delta \hat{\mathbf{d}}_{\mu}(\gamma) \quad (1)$$

where  $\gamma$  is the  $\text{CA}_1\text{-CA}_2\text{-CA}_3$  angle,  $\mathbf{r}_3$  is the position of  $\text{CA}_3$ ,  $\hat{\mathbf{d}}_{\mu}(\gamma)$  is the unit vector indicating the direction of dipole moment, and  $\delta=0.3$  nm.  $\hat{\mathbf{d}}_{\mu}(\gamma)$  is parametrized by fitting on structural data based on a set of non-redundant structures from the PDB-databank, as in ref.;<sup>68</sup> See Supplementary information (SI) for additional details.

**CB reconstruction** The positions of the centroids CB representing the side chains of the amino acids are obtained by a procedure similar to that introduced for the backbone dipoles, and by Levitt.<sup>70</sup> In this case, the unit vector  $\hat{\mathbf{d}}_{\text{CB}}$  pointing at the CB bonded to  $\text{CA}_2$  is expressed as a function of the bending angle  $\gamma$ :

$$\mathbf{r}_{\text{side-chain}} = d_0 \hat{\mathbf{d}}_{\text{CB}}(\gamma) \quad (2)$$

where  $d_0 = 0.38$  nm is the distance to the side chain centroid. In this work, we choose to identify the position of the CB along the atomic  $C_{\alpha} - C_{\beta}$  direction, as previously done in other general models.<sup>70,71</sup> The parametrization was optimized by a weighted least-square fitting on relative orientations of the  $C_{\beta}$  atom gathered from statistics of non-redundant PDB-databank structures (full details on parametrization are given in SI).

## 2.2 hPF potential

In the hPF method, individual polymers are described as a non-interacting systems with hamiltonian  $H_0(\{\mathbf{r}\})$  coupled via a density-dependent interaction functional  $W[\rho(\mathbf{r})]$ :<sup>19,23,24,72–74</sup>

$$H = H_0(\{\mathbf{r}\}) + W[\rho(\mathbf{r})]. \quad (3)$$

For a system made by a number of different particle types, the density dependent functional  $W$  takes the form of:<sup>19</sup>

$$W[\{\rho_n\}] = \frac{1}{\rho_0} \int d\mathbf{r} \left( \frac{k_b T}{2} \sum_{ij} \chi_{ij} \rho_i(\mathbf{r}) \rho_j(\mathbf{r}) + \frac{1}{2\kappa} \left( \sum_j \rho_j - \rho_0 \right)^2 \right). \quad (4)$$

Where  $\rho_i$  is the number density of the  $i^{\text{th}}$  particle type,  $\chi_{ij}$  is the mean field coupling term between particles of type  $i$  and  $j$ ,  $\kappa$  is the compressibility of the system, and  $\rho_0$  is the average number density of the system. The external potential  $V_i^{\text{ext}}(\mathbf{r})$  felt by any individual particle of a generic type  $i$  at position  $\mathbf{r}$  is given by:

$$V_i^{\text{ext}}(\mathbf{r}) = \frac{\delta W[\{\rho_i\}]}{\delta \rho_i(\mathbf{r})} = \frac{1}{\rho_0} \left( k_B T \sum_j \chi_{ij} \rho_j(\mathbf{r}) + \frac{1}{\kappa} \left( \sum_j \rho_j(\mathbf{r}) - \rho_0 \right) \right). \quad (5)$$

### 2.3 Intramolecular potential energy function

The potential energy in hPF model is defined as the sum of the intra-molecular terms, which are based on particle interactions, and the intermolecular ones, determined by the particle density.

Here, the intramolecular term  $V_{\text{intra}}$  takes the form:

$$V_{\text{intra}} = \sum_{\text{bonds}} \frac{1}{2} k_r (r - r_0)^2 + \sum_{\text{bend,torsional}} V_{\gamma,\phi} + \sum_{\text{CA-pairs}} V_{LJ} + \sum_{\text{dipoles-pairs}} V_{\mu_i,\mu_j} \quad (6)$$

the first term is the harmonic stretching potential between two consecutive CA,  $V_{\gamma,\phi}$  is a combined bending and torsional potential acting on the CA chain,  $V_{LJ}$  is a Lennard-Jones potential taking into account the excluded volume between the CA beads along the chain and  $V_{\mu_i,\mu_j}$  is the electrostatic dipole-dipole interaction potential. As a result, all the terms in the potential function are a function of the positions of the CA's, as the positions and orientations of the CB's and the dipoles are both a function of the positions of the CA's.



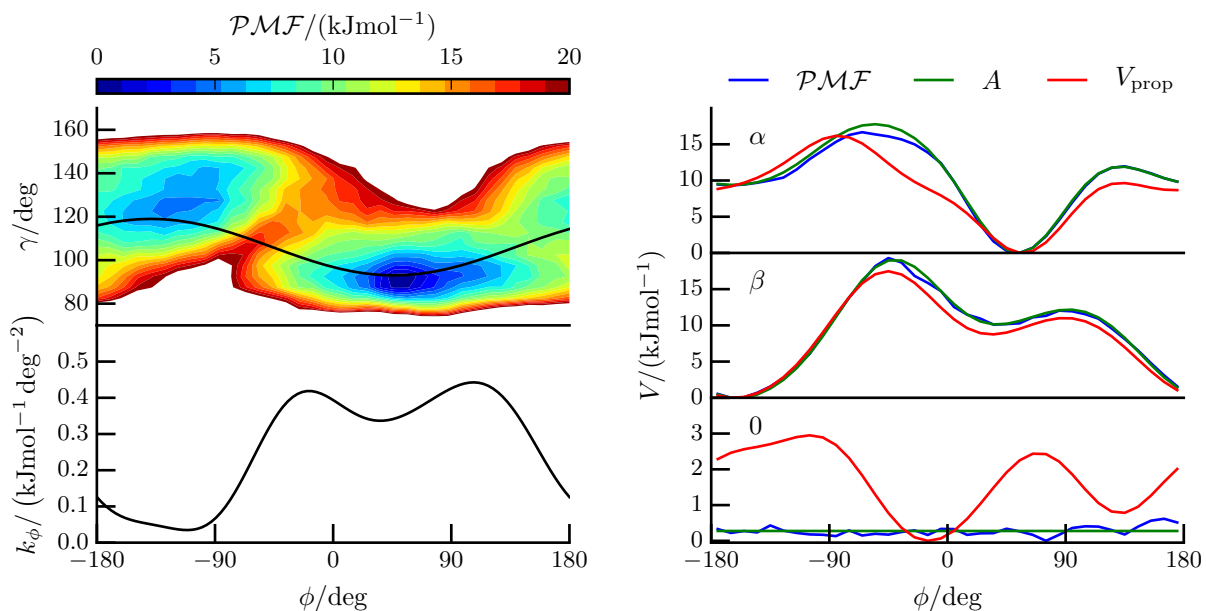


Figure 2: Derivation of the combined bending, torsional potential  $V_{\gamma, \phi}$ . *Top left:* Potential of mean force of an alanine tetrapeptide from all-atom simulations. The  $\gamma_0(\phi)$  line from equation 7 is marked in black. *Bottom left:* Model spring function  $k(\phi)$  in equation 8 used in this work. *Right:* Model propensity potentials  $V_{\text{prop}}(\phi)$  (red lines) for aa's with  $\alpha$ -propensity (top panel),  $\beta$ -propensity (middle-panel), or no propensity (bottom panel), as obtained by iterative Boltzmann inversion. The corresponding potentials of mean force (blue lines) map the ideal target free-energy plots (green lines).

**Bending-torsional potential**  $V_{\gamma,\phi}$  As originally described by Warshel and Levitt,<sup>54</sup> CG bending and torsional angles ( $\gamma$  and  $\phi$  in Figure 1) are strongly correlated, being both order parameters that signal the local coiling of the peptide into different secondary structure elements.<sup>37</sup>

Figure 2 reports the free-energy map for a simple tetra-alanine peptide obtained from metadynamics simulations (See SI 2.1 for technical details). The map evidences the existence of two basins corresponding to helical ( $\phi \simeq 50^\circ$ ) and extended ( $\phi \simeq -180^\circ$ ) conformations. The correlation line proposed by Levitt and Warshel:<sup>54</sup>

$$\gamma_0(\phi) = 106^\circ - 13^\circ \cos(\phi - 45^\circ) \quad (7)$$

falls very close to the variational path connecting the minima of the two basins. It is expected that the qualitative shape of the free energy map is conserved for a generic aa sequence, excluding glycine or proline.<sup>75</sup> Here, we use the following general functional form for  $V_{\gamma,\phi}$ :

$$V_{\gamma,\phi} = V_{\text{prop}}(\phi) + \frac{1}{2}k(\phi)(\gamma - \gamma_0(\phi))^2. \quad (8)$$

where  $k(\phi)$  regulates the harmonic deformations of the chain with respect to  $\gamma_0(\phi)$ , and  $V_{\text{prop}}(\phi)$  is a secondary-structure propensity potential that sculpts the presence and the relative height of any potential energy minimum along  $\phi$ .

**Functional forms of  $k(\phi)$ ,  $V_{\text{prop}}(\phi)$**  Both  $k(\phi)$ , and  $V_{\text{prop}}(\phi)$  functions are expressed as cosine series of the torsional angle:

$$k(\phi) = \sum_{n=0}^4 c_n (1 + \cos(n\phi - d_n)) \quad (9)$$

$$V_{\text{prop}}(\phi) = \sum_{m=0}^4 c_m (1 + \cos(m\phi - d_m)) \quad (10)$$

In the present model, we define  $c_n$ ,  $d_n$  to obtain a model profile for  $k(\phi)$  as in Figure 2, which guarantees the localization of the harmonic fluctuations of the two main basins of stability corresponding to helical and extended conformations, while excluding

unphysical combinations of bending and torsional angles. Values for the coefficients are given in Tables S5-S7.

The coefficients  $c_m$  and the phases  $d_m$  can be optimized to describe the relative thermodynamic stability of local structural features, like coiled versus extended configurations. The propensity toward  $\alpha/\beta$  or other coiling geometries is the result of the free-energy balancing of all the contributions coming from all energy terms, and the entropy. The propensity potential may be calibrated over all atom free energy calculations,<sup>41,76,77</sup> or from experimental data.<sup>78</sup>

For the purposes of this work, we pre-sculpted different model propensities, defined by ideal Helmholtz free-energy profiles  $A(\phi)$ . We then derived the  $V_{\text{prop}}$  yielding a potential of mean force ( $\mathcal{PMF}$ ) close to the target  $A(\phi)$  performing Iterative Boltzmann Inversion.<sup>43</sup> The procedure guarantees excellent agreement for the PMF compared to the target  $A(\phi)$ , with a mean error of 0.4 kJ mol<sup>-1</sup>, and a maximum error of 1.4 kJ mol<sup>-1</sup>, anyway below kT. Technical details of the procedure are given in SI. Figure 2, shows the three model propensities used in this work:  $V_\alpha(\phi)$ ,  $V_\beta(\phi)$ ,  $V_0(\phi)$ , corresponding to aa’s with ideal  $\alpha$ - or  $\beta$ -propensity, or to aa’s with no propensity. Intermediate model propensity potentials were built using the following linear combination:

$$V_{\text{prop}}(\lambda, \phi) = \frac{1}{2} (|\lambda| - \lambda) V_\alpha(\phi) + \frac{1}{2} (|\lambda| + \lambda) V_\beta(\phi) + (1 - |\lambda|) V_0(\phi) \quad (11)$$

where  $\lambda \in [-1, 1]$ .  $\lambda < 0$  values yield  $\alpha$ -propense peptides, and  $\lambda > 0$  values yield  $\beta$ -propense peptides.

## 2.4 Force calculation

Forces acting on dipoles and on CBs are projected onto the CA, whose positions are the only degrees of freedom that are propagated during the molecular dynamics steps.

Specifically,  $F_{ij}^a$ , the  $a^{\text{th}}$  Cartesian component of the force acting on the  $i^{\text{th}}$  CA due

to the  $j^{\text{th}}$  virtual site (either a dipole charge, or a CB), is:

$$F_{ij}^a = -\frac{\partial V(\mathbf{r}_j)}{\partial a_i} = -\frac{\partial V(\mathbf{r}_j)}{\partial \mathbf{r}_j} \frac{\partial \mathbf{r}_j}{\partial a_i} = \sum_{b=x,y,z} F_j^b \frac{\partial b_j}{\partial a_i} \quad (12)$$

where  $\mathbf{r}_j$  is the position of the  $j^{\text{th}}$  virtual site,  $b_j$  is one of its Cartesian components and  $F_j^b$  is the force on site  $j$  in direction  $b$ . Additional details on the derivation of the terms in equation (12) are given in SI.

## 2.5 Simulation details

Table 1: Interaction matrix  $\tilde{\chi} = \chi \cdot RT$  (in  $\text{kJ mol}^{-1}$ ) for an individual aa of the polypeptide chain.  $\alpha \geq 0$  is a control parameter that describes the hydrophobicity of the same aa.

$\tilde{\chi}_{ij}$	CA	CB <sub>P</sub>	CB <sub>H</sub>	H <sub>2</sub> O
CA	0	0	0	0
CB <sub>P</sub>	0	0	$\alpha$	0
CB <sub>H</sub>	0	$\alpha$	0	$\alpha$
H <sub>2</sub> O	0	0	$\alpha$	0

**hPF parameters** The hPF potential controls solute-solute, solute-solvent and solvent-solvent intermolecular interactions. The size of the mesh separation in the grid for the evaluation of the particle-field potential terms was set to 0.65 nm, consistently with the values previously used in simulations of DOPC membranes,<sup>24</sup> where the bond lengths are in the same order of magnitude to that of the peptide model used in this study. The density potential was updated every  $\Delta t = 10 \delta t$ . The compressibility was set to  $\kappa^{-1} = 33.3 \text{ kJ mol}^{-1}$ . The hydrophobic character of each aa was defined by the values of the  $\tilde{\chi}$  matrix  $\tilde{\chi} = \chi \cdot RT$  as in table 1 and equation 5. Specifically, we defined two types of CB's. Type CB<sub>P</sub> is strictly polar, CB<sub>H</sub> is general, and can switch from hydrophilic ( $\alpha = 0$ ) to very hydrophobic aa's ( $\alpha \gg 0$ ). Any value of  $\tilde{\chi}_{ij} = 0$  implies that the two types interact only through the compressibility term.

**Force-field parameters** All systems are described at CG level, with explicit solvent beads. The mass of each CA bead and of the water are all set to 72 amu. The bond-potential has equilibrium distance  $r_0 = 0.38$  nm and spring constant  $k_r = 50000$  kJ nm<sup>-1</sup>. Lennard-Jones parameters  $\sigma = 0.447$  nm and  $\epsilon = 3.5$  kJ mol<sup>-1</sup> are taken from ref.<sup>68</sup> A relative dielectric constant  $\epsilon_r = 5$  is used. Additional details on the evaluation of the electrostatic interactions are given in SI.

**hPF-MD simulations** All simulations are performed in the NVT-ensemble with a density  $\rho_0 = 1.0$  g cm<sup>-3</sup>, at a temperature of 300 K, using a Andersen thermostat with collision frequency  $f_{col} = 7$  ps<sup>-1</sup>. The velocity Verlet algorithm is used to integrate the equations of motion with time-step  $\delta t = 10$  fs. A cutoff  $r_{cut} = 3.5$  nm is used for van der Waals interactions between CA's. Additional force-field parameters are given in SI.

We present three sets of simulations. The first set is on homo-polypeptides of length 15, 30 and 60 aa using type H CB's (CB<sub>H</sub>). The second set of simulations focuses on amphiphilic polypeptides. The third set regards interaction of polypeptides with DOPC bilayer. In these sets, we use mixed H, P types for the CB's. Unless differently stated, in all the simulations the initial geometries of the peptides were set as a random coils. The peptide model was implemented in the OCCAM code, and the same code was used for all the simulations.<sup>72</sup> Details on the setup of the different systems, and the MD parameter sets, are given in SI.

## 3 Results and discussion

### 3.1 Phase diagram of the homopolymer-chain

Within the present model, folding occurs as a balance between the intramolecular potential, driving the local coiling, and the intermolecular forces (hydrophobicity and electrostatics) responsible for the global collapse of the polymer. Figure 3 reports the phase diagrams of homo-polymeric chains as a function of their hydrophobicity and flexibility, obtained by inspecting the statistically relevant stable protein conforma-

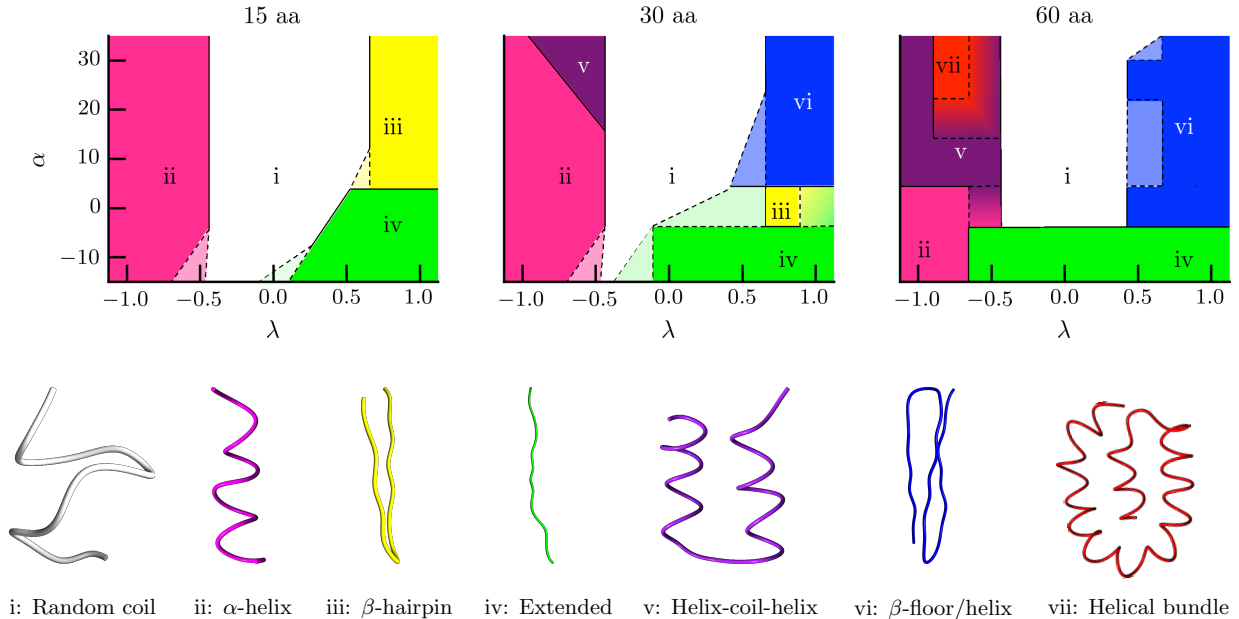
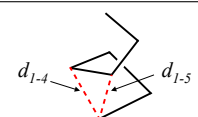
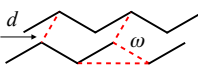


Figure 3: Phase diagrams for homo-peptide sequences. *Top*: Phase diagrams as a function of the coiling propensity parameter  $\lambda$ , and the hydrophobicity parameter  $\alpha$  (in  $\text{kJ mol}^{-1}$ ) for 15 aa, 30 aa and 60 aa long sequences. The regions of stability of different structural motifs are evidenced in different colors and labels. *Bottom*: Structural elements corresponding to the different regions of the phase diagrams.

tions along the MD trajectories. The hydrophobicity was modulated by varying the  $\alpha$  parameter in the  $\tilde{\chi}$  matrix (Table 1) from a value of  $-10 \text{ kJ mol}^{-1}$  (very hydrophilic) to  $30 \text{ kJ mol}^{-1}$  (very hydrophobic). The flexibility was modulated by varying the  $\lambda$  parameter in equation 11.  $\lambda = -1$  describes peptides very prone to coiling,  $\lambda = 0$  ideally flexible peptides, and  $\lambda = 1$  very stiff ones. The phase diagram of short peptides composed by 15 aa is characterized by a simple diagram with four regions of stability corresponding to the  $\alpha$ -helix, the  $\beta$ -hairpin, extended peptide and the random coil configurations (Figure 3a). The structural parameters of the secondary structure elements are presented in Table 2. The 1 – 4 and 1 – 5 CA-distances of helical structures are in excellent agreement with average values of helices from the PDB databank. The  $\beta$ -hairpin inter-strand distance and orientation  $\omega$  are also in very good agreement with experimental values (Table 2). The relatively large fluctuations from the average values have to be attributed to the soft hPF potential compared to particle-particle

distance-dependent ones.

Table 2: Structural parameters for  $\alpha$ -helix and  $\beta$ -structures.

$\alpha$ -helix	$d_{1-4}$ /nm	$d_{1-5}$ /nm	
hPF-MD	$0.51 \pm 0.05$	$0.59 \pm 0.07$	
Ref. (PDB) <sup>68</sup>	$0.51 \pm 0.01$	$0.62 \pm 0.01$	
$\beta$ -hairpin	$d$ /nm	$\omega$ / deg	
hPF-MD	$0.53 \pm 0.04$	$92 \pm 36$	
Ref. (PDB) <sup>68</sup>	$0.48 \pm 0.01$	$90 \pm 6$	

The formation of helical structures is intrinsically determined by the likelihood of peptides to coil. On the contrary, the stabilization of  $\beta$ -hairpins requires a certain degree of hydrophobicity of the chain, which promotes folding by screening from the solvent compensating for the local distortions in the loop region. More hydrophilic peptides are characterized by longer persistence lengths, yielding fully extended structures (Figure 3).

The elongation of the chain to 30 aa produces a complication of the phase diagram, with the appearance of other regions characterized by more complex super-secondary structure assemblies. Similarly for the  $\beta$ -hairpin, the formation of helical super-secondary structures like the helix-coil-helix and super  $\beta$ -structures, is facilitated by more hydrophobic sequences. This is because, with progressively longer sequences, the energy gain due to solvent screening overcomes the costs of introducing a local deformation, for example, at a loop region.

For 60 aa, the length of the polypeptide is beyond the persistence length of individual secondary structure elements in most regions of the diagram. The simple  $\beta$ -hairpin is no longer found, while the single helix is only stable for the most hydrophilic regions and helix-coil-helix has a smaller area of stability.

In all cases, values of  $\lambda \approx 0$ , which corresponds to a flexible ideal-like polymer chain, yield collapsed or elongated structures, while secondary structures appear whenever some preference for either coiling or elongation is introduced. The qualitative features of the phase diagrams here presented are in agreement with previous studies based

on toy models, proposing that flexibility and hydrophobicity, as well as the presence of a directional contact potential, are the three fundamental ingredients required to describe the global folding features of peptides.<sup>79-81</sup>

### 3.2 Amphiphilic peptides

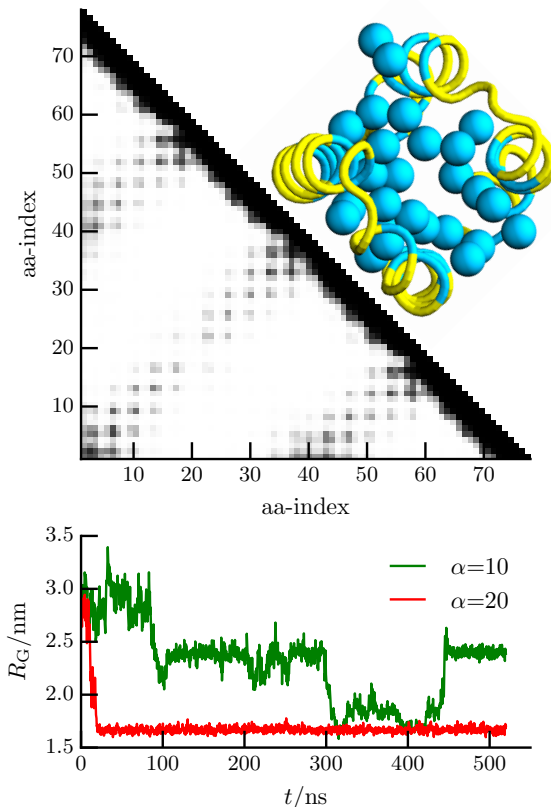


Figure 4: Folding of a four-helix bundle. *Top*: Contact map of the folded structure from hPF-MD simulations, for  $\alpha = 20$  kJ mol<sup>-1</sup>, using a cut-off threshold of 1 nm for the CA-CA distance. The drawing highlights the packing of the hydrophobic CB's in the sequence. Hydrophobic aa's are colored in cyan, hydrophilic ones are colored in yellow. Hydrophilic CB's are not drawn for clarity. *Bottom*: radius of gyration as a function of time for two sequences with different  $\alpha$  values (in kJ mol<sup>-1</sup>).

**Helical peptide – Tertiary-structure assembly** A peptide composed by four segments with strong helical propensity ( $\lambda = -1$ , equation 11) and amphiphilic sequence, linked by short hydrophilic loops with no propensity ( $\lambda = 0$ ) folded locally, as expected,



into four helical moieties. The formation of a compact tertiary assembly was promoted by the collapse of a hydrophobic core, as signaled by the stabilization to a low value (1.7 nm) of the radius of gyration  $R_G$ , when increasing the hydrophobic character from  $\alpha=10$  kJ mol<sup>-1</sup> to  $\alpha=20$  kJ mol<sup>-1</sup> (Figure 4). The folded assembly was stable as a four-helix-bundle, evidenced by the contact map in Figure 4.

The hydrophobic CB's localized in the core of the assembly, where the screening of the solvent is more effective (Figure 4). Solvent screening can be quantified by comparing the water density at the hydrophobic CB's for different values of  $\alpha$  to the bulk water density. The values reported in Table 3 indicate that the stabilization of compact conformations is determined by the tendency of the more hydrophobic CB's to separate from the solvent.

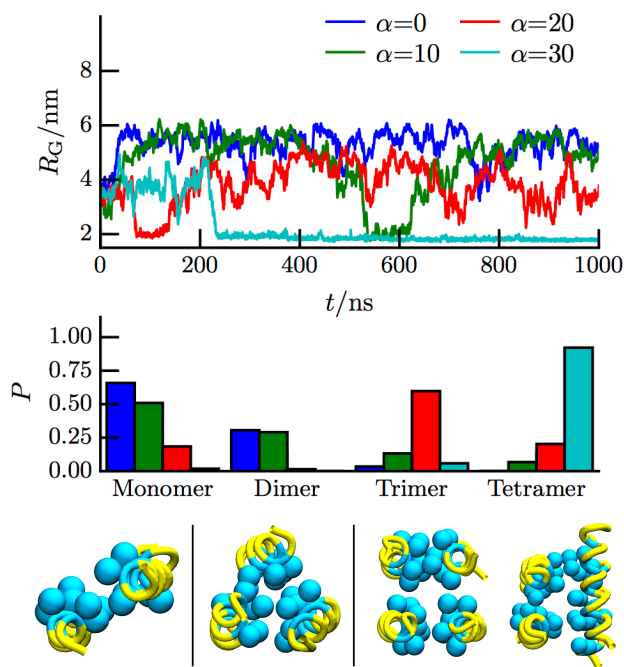


Figure 5: Assembly of four helical amphiphilic peptides. *Top*: Radius of gyration of the four amphiphilic polypeptides for different  $\alpha$  (in kJ mol<sup>-1</sup>) values as a function of time. *Middle*: Relative abundance of aggregates at different  $\alpha$ 's. *Bottom*: relevant structures of the dimeric, trimeric and tetrameric assemblies

**Helical peptide – Quaternary-structure assembly** The same system was simulated removing the loops linking the four helical regions. As evident from Figure

5, also in this case the association of the peptides increased progressively according to the hydrophobic nature of the side-chains. Nonetheless, due to the larger entropic costs for the aggregation of four independent segments, the formation of fully-stable aggregates was observed only for strongly hydrophobic values ( $\alpha = 30 \text{ kJ mol}^{-1}$ ), higher than those necessary in the monomeric sequence. For intermediate values of  $\alpha$ , dimeric and trimeric sequences were observed.

The four-helix complex showed higher conformational disorder than the monomeric peptide, with the co-existence of two quaternary assemblies taking the form of a four (anti)-parallel-helical bundle, or a two+two helical sandwich. This last finding is consistent with former experimental studies on for helix bundles, which demonstrated that folding into one single, well-defined structure can be obtained only in the presence of some chemical variance of the aa sequence.<sup>82-84</sup>

Table 3: Relative water density ( $n_{\text{H}_2\text{O}}$ ) at the hydrophobic CB positions compared to the bulk water density for different values of  $\alpha$  (in  $\text{kJ mol}^{-1}$ ), for the different amphiphilic systems investigated.

Helical peptide		Four helices		$\alpha+\beta$ peptide	
$\alpha$	$n_{\text{H}_2\text{O}}$	$\alpha$	$n_{\text{H}_2\text{O}}$	$\alpha$	$n_{\text{H}_2\text{O}}$
		0	0.45	0	0.42
10	0.27	10	0.34	10	0.29
20	0.11	20	0.23	20	0.20
		30	0.13	30	0.13

**$\alpha+\beta$ -peptide.** Starting from a random-coil conformation, the two regions of the polypeptide folded into their respective local secondary structures. Nonetheless, for hydrophilic sequences, no tertiary organization was observed as indicated by the corresponding featureless contact map (Figure 6). The increase in the hydrophobic parameter value was accompanied to the appearance of more pronounced contact maps, indicating the formation of more compact structures and more resilient contacts. Like in the helical peptide, the formation of a clearly folded tertiary complex was found for a value of  $\alpha=20 \text{ kJ mol}^{-1}$ , where a single helix packs over a  $\beta$ -hairpin (Figure 6).

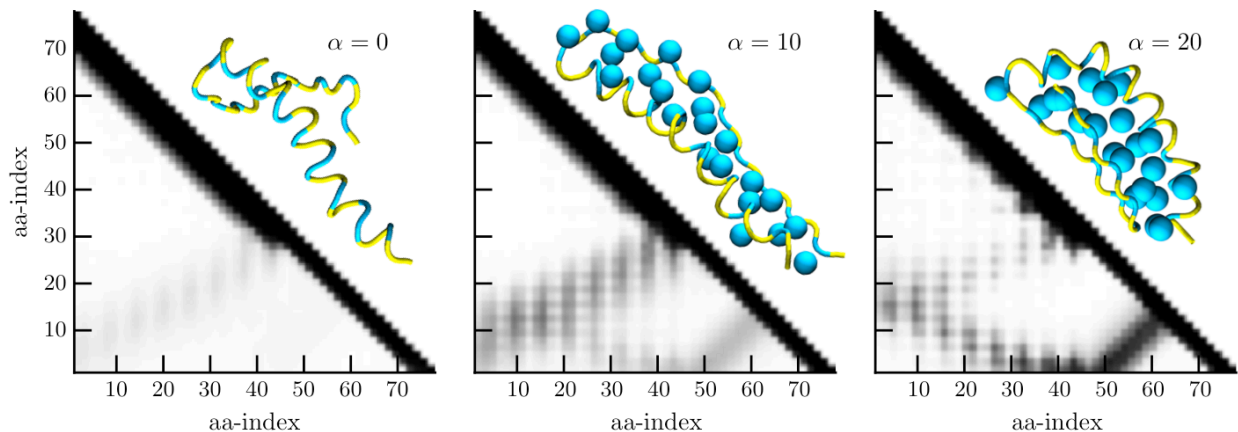


Figure 6: Folding of  $\alpha+\beta$  peptide. The three panels report the contact maps for different values of  $\alpha = 0, 10, 20$  kJ mol<sup>-1</sup>, and the corresponding most abundant structures. Hydrophobic aa's are colored in cyan, hydrophilic ones are colored in yellow. Hydrophilic CB's are not drawn for clarity.

At intermediate hydrophobicity ( $\alpha=10$  kJ mol<sup>-1</sup>), we observed the existence of a molten globule-like structure, where the properly folded compact assembly was in competition with a more solvent exposed assembly formed by a single helix packed to a single  $\beta$ -strand (Figure 6). The appearance of assemblies of individual structure elements characterized by longer persistence lengths is consistent with what observed in the the phase diagram for homo-polymer sequences.

### 3.3 Interaction of peptides with a lipid bilayer

**Hydrophobic peptide** We considered two systems consisting of a 25 aa long polypeptide with hydrophilic ends composed each by 3 aa and a hydrophobic ( $\alpha = 30$  kJ mol<sup>-1</sup>) central region. A weak or a medium helical propensity ( $\lambda = -0.25, -0.50$  respectively) was attributed to either of the two sequences. Both peptides were either fully solvated in water, or embedded into a DOPC bilayer. Consistently with the phase diagrams of homopolymer sequences, the weakly helical peptide showed no structure in water, while the peptide with more helical propensity folded into a helix-coil-helix (Figure 7). These results show that, when in water, the hydrophobic region determines the folding of the peptide, while the short hydrophilic ends do not affect the behavior

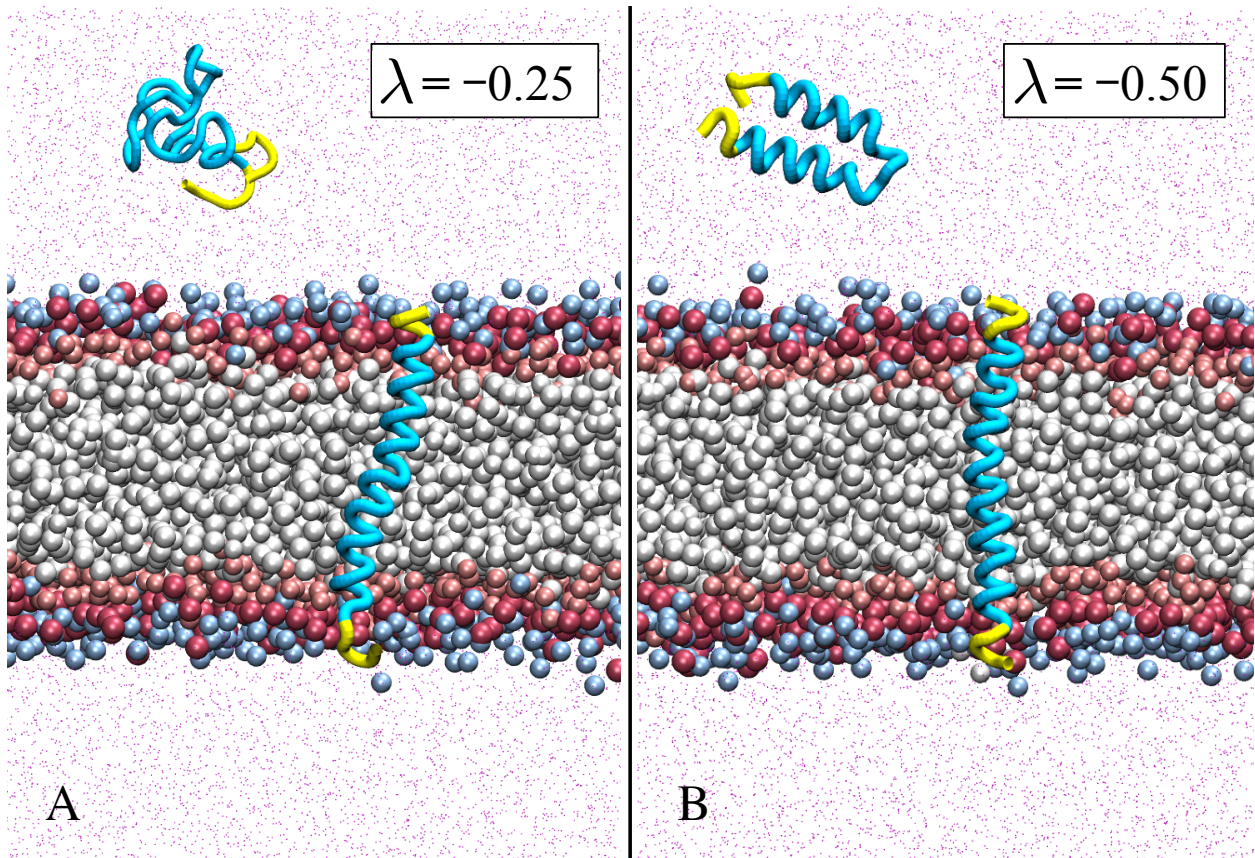


Figure 7: Interaction of an hydrophobic peptide with a DOPC bilayer. The left and right panels report the structures in water and in the membrane of two peptides having the propensity parameter  $\lambda = -0.25$  and  $\lambda = -0.5$ , respectively. Hydrophobic residues are colored in cyan, the hydrophilic ends are colored in yellow. The DOPC bilayer is represented in spheres. Water beads are represented by dots.

of the peptide.

When the same peptides were inserted into a DOPC bilayer, we instead observed a stable single helix over the 500 ns of MD (Figure 7). Interestingly, such conditions helped the stabilization into a helical structure for a propensity value as low as ( $\lambda = -0.25$ ), for which the phase diagram in the homogeneous aqueous phase predicts no stable helices regardless of the sequence hydrophobicity. The trans-membrane helical folding is stabilized by the localization of the hydrophilic ends at the exterior of the bilayer, while keeping the hydrophobic region of the peptide into the lipid core. In fact, the length of the hydrophobic helical segment (3.8(2) nm) is similar to the distance between the glycerol bead planes in the two leaflets (3.22(1) nm), which define the

hydrophobic region of the DOPC bilayer.

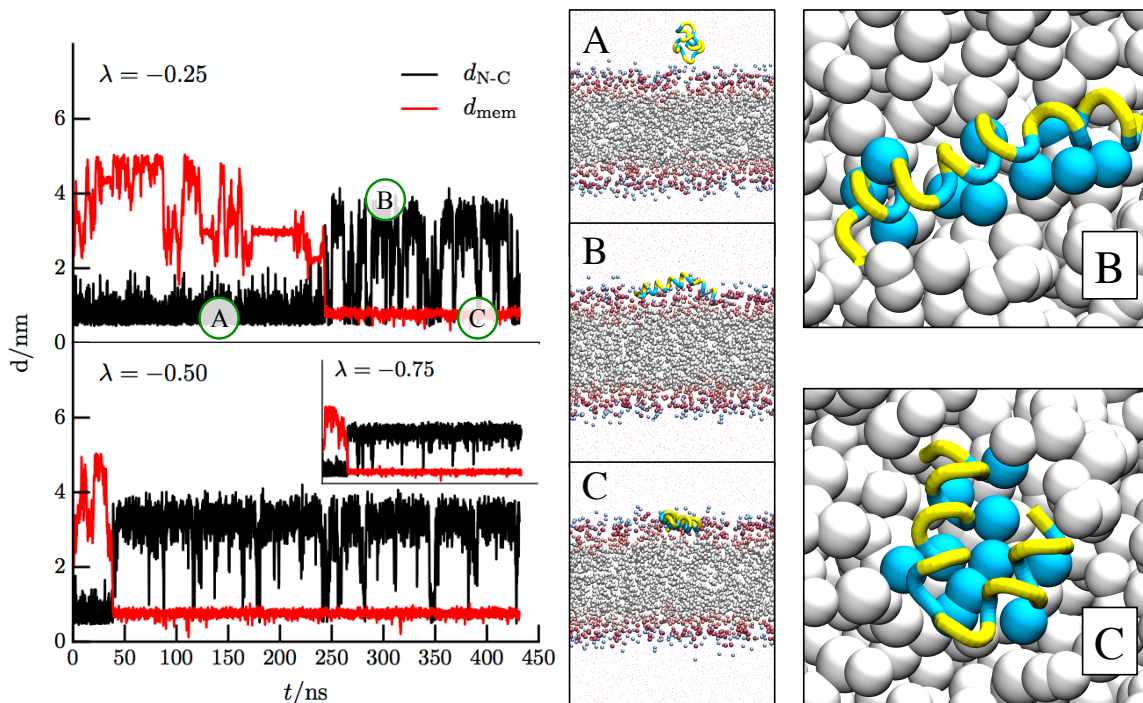


Figure 8: Interaction of an amphiphilic peptide with a DOPC bilayer. *Left*: the plots report the distance of the center of mass of the peptide from the DOPC membrane plane, defined by the G beads of the lipids ( $d_{\text{mem}}$ , red line), and the N-to-C distance of the peptide ( $d_{\text{N-C}}$ , black line) as a function of time for three different peptides with different propensity parameter  $\lambda$ . the A, B, C letters mark three snapshot taken for the simulation with  $\lambda=-0.25$ . *Center*: Side-view of the three A, B, C geometries. Hydrophobic residues are colored in cyan, the hydrophilic ones are colored in yellow. The DOPC bilayer is represented in spheres. Water beads are represented by dots. *Right*: Top-view of the B, C snapshots. Hydrophobic CB's are drawn in cyan spheres, DOPC beads are in white spheres. The hydrophilic CB's as well as water are not shown for clarity.

**Amphiphilic peptide** A peptide with amphiphilic helical character was placed in the aqueous phase in the neighborhood of the DOPC bilayer. Three simulations were repeated varying the helical propensity from low ( $\lambda = -0.25$ ) to medium ( $\lambda = -0.50$ ), to strong ( $\lambda = -0.75$ ). In water, the peptides showed tendency to form compact structures, either as a collapsed coil ( $\lambda = -0.25$ , Figure 8, A panels), or as helix-coil-helix ( $\lambda = -0.50, -0.75$ ), screening the hydrophobic side-chains from the solvent.

Figure 8 reports both the end-to-end distance, and the distance of the center of mass of the peptide from the membrane plane, defined by the glycerol beads in the

lipid monolayer closest to the peptide. In all cases, the peptides bound to the membrane surface. The localization of the peptide was quantified by computing the density profile along the axis normal to the membrane. The plot, reported in figure S3, shows that the peptide lies in the region of the polar heads, with the hydrophobic CB's pointing toward the glycerol beads, and the hydrophilic CB's remaining exposed to the solvent.

Binding to the membrane was associated to a rapid folding of all the peptides in a helical structure, as indicated by the transition to high values of the end-to-end distance (Figure 8). Interestingly, the interaction with the membrane promoted helical features also for  $\lambda$  values smaller than the coil/helix transition threshold in the homogeneous phase. Transversal deformations of the helix may be promoted either by fluctuations of the membrane, or by the conformational entropy of the peptide peptide, leave the peptide attached to the membrane plane, and thus do not affect the screening of the hydrophobic chains. In fact, all helix-coil-helix structure, in competition with the single-helix, were observed (Figure 8, panels B, C). for the three  $\lambda = -0.25, -0.50, -0.75$  values, we found helix-coil-helix:single-helix probability ratios of 1:1.3, 1:6.8 and 1:18.1, respectively, showing how the relative stability of single helix is promoted by a stronger helical propensity of the sequence.

## 4 Concluding remarks

We developed and implemented a hydrophobic-polar model for polypeptides based on a two-bead mapping of the amino acids, where the intermolecular interactions are described by density-functional based hPF potentials. We simulated a series of test systems verifying that the model here introduced reproduces the main structural and dynamical features of polypeptide chains. In particular, we could observe: *(i)* the existence of separate regions of stability for folded, collapsed, and swollen structures as a function of the flexibility and hydrophobicity of the sequence; *(ii)* the effect of the hydrophobic core in driving the assembling of simple tertiary and quaternary structures, and *(iii)* that sequences with mixed hydrophobic/hydrophilic patterns localize

and fold differently at water/lipid interfaces than in the bulk solvent.

The functional form for the peptide backbone ensures the presence of structural minima for the canonical helical or extended conformations. At the same time, the peptide chain is able to respond to environmental perturbations and to assume locally distorted conformations, whenever the enthalpy/entropy balance is globally favorable. With this respect, the use of an explicit solvent model guarantees that solvent release is properly taken into account during chain-packing events. Our model showed good capability of responding to environmental changes, like the presence of a lipid bilayer, both by predicting the proper localization of the peptide, and by being able to predict its membrane binding-dependent refolding.

Our model uses a topological reconstruction of the peptide electrostatic dipoles as introduced in previous publications.<sup>67–69</sup> These elements guarantee both the stabilization of folded structures by mimicking H-bond formation, and a sound description of long-range electrostatics,<sup>67</sup> which may be a crucial ingredient for future studies on macromolecular crowding. Given the low costs and high parallelization of the hPF method,<sup>24</sup> the computational efficiency of the present model is critically dependent on the costs for the electrostatic term. The recent development of an electrostatic-consistent hPF scheme<sup>85</sup> ensures that the presence of explicit electrostatic terms in the molecular Hamiltonian do not become a serious bottleneck for large-scale applications.<sup>29</sup> Future assessment on the performance of hPF with electrostatics and related developments on the same topic may be of greater help for hPF simulations of biological systems, especially thinking to possible extension of the model to nucleic acids.

Finally, here we limited our exploration to model hydrophobic/hydrophilic peptide sequences. Moreover, we have introduced ideal  $\alpha/\beta$  propense potentials that do not correspond to specific amino acid. Thus, the range of applicability of the model, at present, limits to general biophysical studies, and should not be used to describe specific biochemical sequences. Nonetheless, we remark that the bending-torsional potential, and in particular, its propensity term  $V_{\text{prop}}(\phi)$ , can be calibrated to quantitatively reproduce the secondary structure propensities of any specific amino acid. For this

reason, the present model represents an important first step toward the establishment of a new hPF-based CG force-field for proteins able to describe structural and dynamical features with biochemical detail.

## Acknowledgement

Authors acknowledge the support of the Norwegian Research Council through the CoE Hylleraas Centre for Quantum Molecular Sciences (Grant n. 262695), and by the Norwegian Supercomputing Program (NOTUR) (Grant No. NN4654K). Authors thank Franca Fraternali for giving critical comments on the manuscript.

**Associated Content** Protocol for the reconstruction of the dipoles and the CB, equations for the force distribution, details on the derivation of the bending-torsional potential, simulation details, density profile of the amphiphilic peptide bound to a DOPC bilayer.

## References

- (1) Kawakatsu, T. *Statistical physics of polymers: an introduction*; Springer, 2013.
- (2) Edwards, S. F. The statistical mechanics of polymers with excluded volume. *Proc. Phys. Soc.* **1965**, *85*, 613–624.
- (3) Helfand, E.; Tagami, Y. Theory of the interface between immiscible polymers. *J. Polym. Sci. Pol. Lett.* **1971**, *9*, 741–746.
- (4) Helfand, E.; Tagami, Y. Theory of the interface between immiscible polymers. II. *J. Chem. Phys.* **1972**, *57*, 1812–1813.
- (5) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.



- (6) Milner, S. T.; Witten, T.; Cates, M. Theory of the grafted polymer brush. *Macromolecules* **1988**, *21*, 2610–2619.
- (7) Netz, R. R.; Schick, M. Polymer brushes: from self-consistent field theory to classical theory. *Macromolecules* **1998**, *31*, 5105–5122.
- (8) Balazs, A. C.; Singh, C.; Zhulina, E. Modeling the Interactions between Polymers and Clay Surfaces through Self-Consistent Field Theory. *Macromolecules* **1998**, *31*, 8370–8381.
- (9) Matsen, M. W.; Schick, M. Stable and unstable phases of a diblock copolymer melt. *Phys. Rev. Lett.* **1994**, *72*, 2660–2663.
- (10) Drolet, F.; Fredrickson, G. H. Combinatorial screening of complex block copolymer assembly with self-consistent field theory. *Phys. Rev. Lett.* **1999**, *83*, 4317–4320.
- (11) Roan, J.-R.; Kawakatsu, T. Self-consistent-field theory for interacting polymeric assemblies. I. Formulation, implementation, and benchmark tests. *J. Chem. Phys.* **2002**, *116*, 7283–7294.
- (12) Daoulas, K. C.; Müller, M.; De Pablo, J. J.; Nealey, P. F.; Smith, G. D. Morphology of multi-component polymer systems: single chain in mean field simulation studies. *Soft Matter* **2006**, *2*, 573–583.
- (13) Müller, M.; Smith, G. D. Phase separation in binary mixtures containing polymers: A quantitative comparison of single-chain-in-mean-field simulations and computer simulations of the corresponding multichain systems. *J. Polym. Sci. Pol. Phys.* **2005**, *43*, 934–958.
- (14) Stoykovich, M. P.; Müller, M.; Kim, S. O.; Solak, H. H.; Edwards, E. W.; De Pablo, J. J.; Nealey, P. F. Directed assembly of block copolymer blends into nonregular device-oriented structures. *Science* **2005**, *308*, 1442–1446.

- (15) Detcheverry, F. A.; Kang, H.; Daoulas, K. C.; Müller, M.; Nealey, P. F.; de Pablo, J. J. Monte Carlo simulations of a coarse grain model for block copolymers and nanocomposites. *Macromolecules* **2008**, *41*, 4989–5001.
- (16) Kang, H.; Detcheverry, F. A.; Mangham, A. N.; Stoykovich, M. P.; Daoulas, K. C.; Hamers, R. J.; Müller, M.; de Pablo, J. J.; Nealey, P. F. Hierarchical assembly of nanoparticle superstructures from block copolymer-nanoparticle composites. *Phys. Rev. Lett.* **2008**, *100*, 148303–4.
- (17) Behringer, H.; Qi, S.; Schmid, F. Using field theory to construct hybrid particle-continuum simulation schemes with adaptive resolution for soft matter systems. *New J. Phys.* **2013**, *15*, 125009.
- (18) Qi, S.; Behringer, H.; Raasch, T.; Schmid, F. A hybrid particle-continuum resolution method and its application to a homopolymer solution. *Eur. Phys. J. - Spec. Top.* **2016**, *225*, 1527–1549.
- (19) Milano, G.; Kawakatsu, T. Hybrid particle-field molecular dynamics simulations for dense polymer systems. *J. Chem. Phys.* **2009**, *130*, 214106–8.
- (20) Zhao, Y.; De Nicola, A.; Kawakatsu, T.; Milano, G. Hybrid particle-field molecular dynamics simulations: Parallelization and benchmarks. *J. Comput. Chem.* **2012**, *33*, 868–880.
- (21) Milano, G.; Kawakatsu, T. Pressure calculation in hybrid particle-field simulations. *J. Chem. Phys.* **2010**, *133*, 214102–13.
- (22) De Nicola, A.; Kawakatsu, T.; Milano, G. Generation of Well-Relaxed All-Atom Models of Large Molecular Weight Polymer Melts: A Hybrid Particle-Continuum Approach Based on Particle-Field Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 5651–5667.
- (23) De Nicola, A.; Zhao, Y.; Kawakatsu, T.; Roccatano, D.; Milano, G. Hybrid

- particle-field coarse-grained models for biological phospholipids. *J. Chem. Theory Comput.* **2011**, *7*, 2947–2962.
- (24) Milano, G.; Kawakatsu, T.; De Nicola, A. A hybrid particle–field molecular dynamics approach: a route toward efficient coarse-grained models for biomembranes. *Phys. Biol.* **2013**, *10*, 045007–16.
- (25) De Nicola, A.; Kawakatsu, T.; Rosano, C.; Celino, M.; Rocco, M.; Milano, G. Self-assembly of Triton X-100 in water solutions: a multiscale simulation study linking mesoscale to atomistic models. *J. Chem. Theory Comput.* **2015**, *11*, 4959–4971.
- (26) Sevink, G. J. A.; Schmid, F.; Kawakatsu, T.; Milano, G. Adding cell-based hydrodynamics to hybrid particle-field simulations: Efficient and realistic simulation of structuring dynamics. *Soft Matter* **2017**, *13*, 1594–1623.
- (27) De Nicola, A.; Hezaveh, S.; Zhao, Y.; Kawakatsu, T.; Roccatano, D.; Milano, G. Micellar drug nanocarriers and biomembranes: how do they interact? *Phys. Chem. Chem. Phys.* **2014**, *16*, 5093–5105.
- (28) Pizzirusso, A.; De Nicola, A.; Sevink, G. J. A.; Correa, A.; Cascella, M.; Kawakatsu, T.; Rocco, M.; Zhao, Y.; Celino, M.; Milano, G. Biomembrane solubilization mechanism by Triton X-100: a computational study of the three stage model. *Phys. Chem. Chem. Phys.* **2017**, *19*, 29780–29794.
- (29) Soares, T. A.; Vanni, S.; Milano, G.; Cascella, M. Toward Chemically Resolved Computer Simulations of Dynamics and Remodeling of Biological Membranes. *J. Phys. Chem. Lett.* **2017**, *8*, 3586–3594.
- (30) Wang, K.; McClure, J.; Tu, A. Titin: major myofibrillar components of striated muscle. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 3698–3702.
- (31) Perumal, S.; Antipova, O.; Orgel, J. P. Collagen fibril architecture, domain organization, and triple-helical conformation govern its proteolysis. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2824–2829.

- (32) Holmes, K. C.; Popp, D.; Gebhard, W.; Kabsch, W. Atomic model of the actin filament. *Nature* **1990**, *347*, 44–49.
- (33) Salditt, T.; Li, C.; Spaar, A. Structure of antimicrobial peptides and lipid membranes probed by interface-sensitive X-ray scattering. *BBA-Biomembranes* **2006**, *1758*, 1483–1498.
- (34) Ponder, J. W.; Case, D. A. *Protein Simulations*; Advances in Protein Chemistry Supplement C; Academic Press, 2003; Vol. 66; pp 27–85.
- (35) Tringe, J. W.; Ileri, N.; Levie, H. W.; Stroeve, P.; Ustach, V.; Faller, R.; Renaud, P. Molecular Dynamics and Monte Carlo simulations resolve apparent diffusion rate differences for proteins confined in nanochannels. *Chem. Phys.* **2015**, *457*, 19–27.
- (36) Tozzini, V. Coarse-grained models for proteins. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144–150.
- (37) Tozzini, V. Minimalist models for proteins: a comparative analysis. *Q. Rev. Biophys.* **2010**, *43*, 333–371.
- (38) Cascella, M.; Vanni, S. *Chemical Modelling*; The Royal Society of Chemistry, 2016; Vol. 12; pp 1–52.
- (39) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. V. Coarse-Grained protein models and their applications. *Chem. Rev.* **2016**, *116*, 7898–7936.
- (40) Ayton, G. S.; Noid, W. G.; Voth, G. A. Multiscale modeling of biomolecular systems: in serial and in parallel. *Curr. Opin. Struct. Biol.* **2007**, *17*, 192–198.
- (41) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819–834.

- (42) DeVane, R.; Shinoda, W.; Moore, P. B.; Klein, M. L. Transferable coarse grain nonbonded interaction model for amino acids. *J. Chem. Theory Comput.* **2009**, *5*, 2115–2124.
- (43) Reith, D.; Pütz, M.; Müller-Plathe, F. Deriving effective mesoscale potentials from atomistic simulations. *J. Comput. Chem.* **2003**, *24*, 1624–1636.
- (44) Ercolessi, F.; Adams, J. B. Interatomic Potentials from First-Principles Calculations: The Force-Matching Method. *Eur. Phys. Lett.* **1994**, *26*, 583–588.
- (45) Izvekov, S.; Voth, G. A. A multiscale coarse-graining method for biomolecular systems. *J. Phys. Chem. B* **2005**, *109*, 2469–2473.
- (46) Noid, W.; Chu, J.-W.; Ayton, G. S.; Krishna, V.; Izvekov, S.; Voth, G. A.; Das, A.; Andersen, H. C. The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. *J. Chem. Phys.* **2008**, *128*, 244114–11.
- (47) Voth, G. A. *Coarse-graining of condensed phase and biomolecular systems*; CRC press, 2008.
- (48) Zhou, J.; Thorpe, I. F.; Izvekov, S.; Voth, G. A. Coarse-grained peptide modeling using a systematic multiscale approach. *Biophys. J.* **2007**, *92*, 4289–4303.
- (49) Grime, J. M.; Voth, G. A. Early stages of the HIV-1 capsid protein lattice formation. *Biophys. J.* **2012**, *103*, 1774–1783.
- (50) Chu, J.-W.; Voth, G. A. Coarse-grained modeling of the actin filament derived from atomistic-scale simulations. *Biophys. J.* **2006**, *90*, 1572–1582.
- (51) Davtyan, A.; Simunovic, M.; Voth, G. A. Multiscale simulations of protein-facilitated membrane remodeling. *J. Struct. Biol.* **2016**, *196*, 57–63.

- (52) Mim, C.; Cui, H.; Gawronski-Salerno, J.; Frost, A.; Lyman, E.; Voth, G.; Unger, V. Structural Basis of Membrane Bending by the N-BAR Protein Endophilin. *Cell* **2012**, *149*, 137–145.
- (53) Liwo, A.; Ołdziej, S.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. A united-residue force field for off-lattice protein-structure simulations. I. Functional forms and parameters of long-range side-chain interaction potentials from protein crystal data. *J. Comput. Chem.* **1997**, *18*, 849–873.
- (54) Levitt, M.; Warshel, A. Computer simulation of protein folding. *Nature* **1975**, *253*, 694–698.
- (55) Carmichael, S. P.; Shell, M. S. A New Multiscale Algorithm and Its Application to Coarse-Grained Peptide Models for Self-Assembly. *J. Phys. Chem. B* **2012**, *116*, 8383–8393.
- (56) Sagui, C.; Darden, T. A. Molecular dynamics simulations of biomolecules: long-range electrostatic effects. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155–179.
- (57) Warshel, A.; Russell, S. T. Calculations of electrostatic interactions in biological systems and in solutions. *Q. Rev. Biophys.* **1984**, *17*, 283–422.
- (58) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **2013**, *9*, 687–697.
- (59) Basdevant, N.; Borgis, D.; Ha-Duong, T. A Coarse-Grained Protein-Protein Potential Derived from an All-Atom Force Field. *J. Phys. Chem. B* **2007**, *111*, 9390–9399.
- (60) Terakawa, T.; Takada, S. RESPAC: Method to Determine Partial Charges in

- Coarse-Grained Protein Model and Its Application to DNA-Binding Proteins. *J. Chem. Theory Comput.* **2014**, *10*, 711–721.
- (61) Basdevant, N.; Borgis, D.; Ha-Duong, T. Modeling Protein-Protein Recognition in Solution Using the Coarse-Grained Force Field SCORPION. *J. Chem. Theory Comput.* **2013**, *9*, 803–813.
- (62) Irbäck, A.; Sjunnesson, F.; Wallin, S. Three-helix-bundle protein in a Ramachandran model. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13614–13618.
- (63) Bereau, T.; Deserno, M. Generic coarse-grained model for protein folding and aggregation. *J. Chem. Phys.* **2009**, *130*, 235106–15.
- (64) Messer, B. M.; Roca, M.; Chu, Z. T.; Vicatos, S.; Kilshtain, A. V.; Warshel, A. Multiscale simulations of protein landscapes: Using coarse-grained models as reference potentials to full explicit models. *Proteins* **2010**, *78*, 1212–1227.
- (65) Kar, P.; Gopal, S. M.; Cheng, Y.-M.; Predeus, A.; Feig, M. PRIMO: a transferable coarse-grained force field for proteins. *J. Chem. Theory Comput.* **2013**, *9*, 3769–3788.
- (66) Darr, L.; Machado, M. R.; Brandner, A. F.; Gonzalez, H. C.; Ferreira, S.; Pantano, S. SIRAH: A Structurally Unbiased Coarse-Grained Force Field for Proteins with Aqueous Solvation and Long-Range Electrostatics. *J. Chem. Theory Comput.* **2015**, *11*, 723–739.
- (67) Cascella, M.; Neri, M. A.; Carloni, P.; Dal Peraro, M. Topologically based multipolar reconstruction of electrostatic interactions in multiscale simulations of proteins. *J. Chem. Theory Comput.* **2008**, *4*, 1378–1385.
- (68) Alemani, D.; Collu, F.; Cascella, M.; Dal Peraro, M. A nonradial coarse-grained potential for proteins produces naturally stable secondary structure elements. *J. Chem. Theory Comput.* **2010**, *6*, 315–324.

- (69) Spiga, E.; Alemani, D.; Degiacomi, M. T.; Cascella, M.; Dal Peraro, M. Electrostatic-Consistent Coarse-Grained Potentials for Molecular Simulations of Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 3515–3526.
- (70) Park, B.; Levitt, M. Energy Functions that Discriminate X-ray and Near-native Folds from Well-constructed Decoys. *J. Mol. Biol.* **1996**, *258*, 367–392.
- (71) Micheletti, C.; Carloni, P.; Maritan, A. Accurate and efficient description of protein vibrational dynamics: comparing molecular dynamics and Gaussian models. *Proteins* **2004**, *55*, 635–645.
- (72) Zhao, Y.; De Nicola, A.; Kawakatsu, T.; Milano, G. Hybrid particle-field molecular dynamics simulations: Parallelization and benchmarks. *J. Comput. Chem.* **2012**, *33*, 868–880.
- (73) De Nicola, A.; Kawakatsu, T.; Milano, G. A Hybrid Particle-Field Coarse-Grained Molecular Model for Pluronics Water Mixtures. *Macromol. Chem. Phys.* **2013**, *214*, 1940–1950.
- (74) Sarukhanyan, E.; De Nicola, A.; Roccatano, D.; Kawakatsu, T.; Milano, G. Spontaneous insertion of carbon nanotube bundles inside biomembranes: a hybrid particle-field coarse-grained molecular dynamics study. *Chem. Phys. Lett.* **2014**, *595*, 156–166.
- (75) Ha-Duong, T. Protein Backbone Dynamics Simulations Using Coarse-Grained Bonded Potentials and Simplified Hydrogen Bonds. *J. Chem. Theory Comput.* **2010**, *6*, 761–773.
- (76) Bahar, I.; Jernigan, R. L. Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. *J. Mol. Biol.* **1997**, *266*, 195–214.
- (77) Tozzini, V.; McCammon, J. A. *Coarse-graining of condensed Phase and Biomolecular Systems*; CRC Press, 2008; pp 285–298.



- (78) Tozzini, V.; Rocchia, W.; McCammon, J. A. Mapping All-Atom Models onto One-Bead Coarse-Grained Models: General Properties and Applications to a Minimal Polypeptide Model. *J. Chem. Theory Comput.* **2006**, *2*, 667–673.
- (79) Hoang, T. X.; Trovato, A.; Seno, F.; Banavar, J. R.; Maritan, A. Geometry and symmetry prescript the free-energy landscape of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7960–7964.
- (80) Chen, N.-Y.; Su, Z.-Y.; Mou, C.-Y. Effective potentials for folding proteins. *Phys. Rev. Lett.* **2006**, *96*, 078103.
- (81) Skrbčić, T.; Hoang, T. X.; Giacometti, A. Effective stiffness and formation of secondary structures in a protein-like model. *J. Chem. Phys.* **2016**, *145*, 084904–10.
- (82) Choma, C. T.; Lear, J. D.; Nelson, M. J.; Dutton, P. L.; Robertson, D. E.; De Grado, W. F. Design of a heme-binding four-helix bundle. *J. Am. Chem. Soc.* **1994**, *116*, 856–865.
- (83) Kamtekar, S.; Hecht, M. H. Protein Motifs. 7. The four-helix bundle: what determines a fold? *FASEB J.* **1995**, *9*, 1013–1022.
- (84) Hecht, M.; Richardson, J.; Richardson, D.; Ogden, R. De novo design, expression, and characterization of Felix: a four-helix bundle protein of native-like sequence. *Science* **1990**, *249*, 884–891.
- (85) Zhu, Y.-L.; Lu, Z.-Y.; Milano, G.; Shi, A.-C.; Sun, Z.-Y. Hybrid particle–field molecular dynamics simulation for polyelectrolyte systems. *Phys. Chem. Chem. Phys.* **2016**, *18*, 9799–9808.

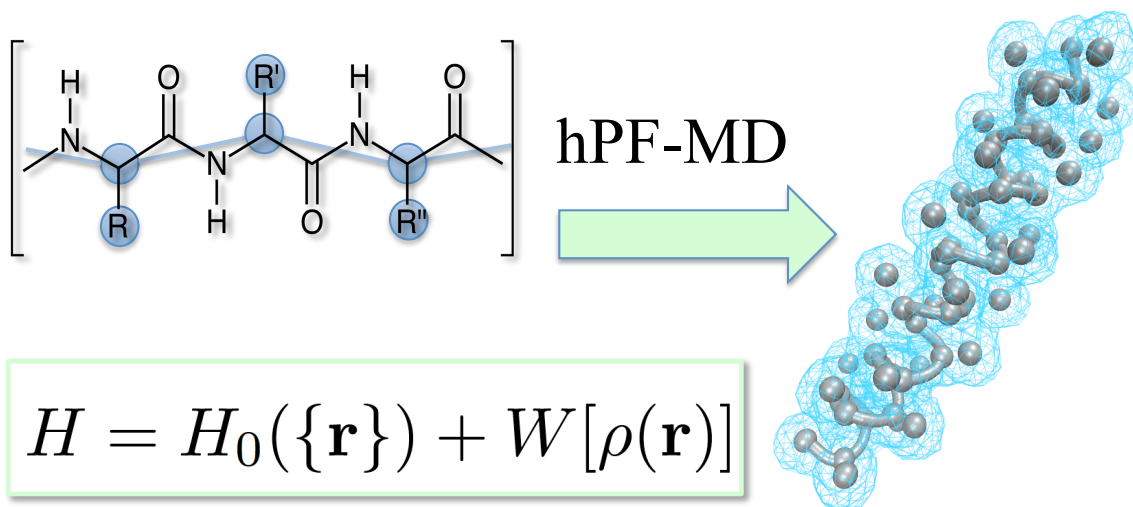


Table of Contents.