Room for Improvement in the Initial Martini 3 Parameterization of Peptide Interactions

J. Karl Spinti, Fernando Neiva Nunes, Manuel N. Melo*

Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Avenida da República, 2780-157, Oeiras, Portugal

Abstract

The Martini 3 coarse-grain force field has greatly improved upon its predecessor, having already been successfully employed in several applications. Here, we gauge the accuracy of Martini 2 and 3 protein interactions in two types of systems: coiled coil peptide dimers in water and transmembrane peptides. Coiled coil dimers form incorrectly under Martini 2 and not at all under Martini 3. With transmembrane peptides, Martini 3 represents better the membrane thickness–peptide tilt relationship, but shorter peptides do not remain transmembranar. We discuss related observations, and describe mitigation strategies involving either scaling interactions or restraining the system.

Keywords: Coarse-grain, Molecular dynamics, Martini, Coiled coil, Transmembrane, Peptides

1. Introduction

The past decade has seen an explosion in the adoption of coarse-grain (CG) molecular dynamics (MD) [1]. By employing reduced-dimensionality models one can lower the computational costs of MD simulations by several orders of magnitude, and bring into reach timescales of milliseconds and size scales approaching the micrometer [1]. The Martini force field has been at the forefront of many important CG MD applications [1] and is arguably the most used such model for biomolecular simulations. This is not to say CG models like Martini — where simplification is achieved by representing groups of about 4 non-hydrogen atoms as single particles, or “beads” — are without limitations: besides obvious trade-offs, such as the inability to represent fine chirality or explicit hydrogen-bonding, Martini often suffered from “overmapping”, a condition in which models with mappings finer that 4-to-1 led to imbalanced hydrophobic/hydrophilic interactions [2].

The recently-released Martini 3 model [3] tackles the limitations of the prior Martini 2 model [4–6] by introducing several new particle types, with explicitly different interaction sizes to account for 3-to-1 and 2-to-1 mappings. With this new approach proteins in solution, transmembrane (TM) pores, and helical membrane-bound oligomers have been successfully simulated without undergoing excessive protein aggregation [3] as was likely under Martini 2 [2, 7–9]. With over-aggregation out of the way, we set out in this work to gauge the accuracy of more subtle aspects of Martini 3 protein interactions. We focus on two types of biologically relevant systems: i) coiled-coil dimers with well-defined parallel structures [10, 11] (PDB ids 1ZIK and 1U0I) to test Martini 3 recognition of heptad repeats – an aspect strongly dependent on steric matching and hydrophobicity/hydrophilicity discrimination [12]; and ii) the WALP (tryptophan–alanine–leucine) series of peptides, which are single-pass TM alpha-helices with a well described tilting behavior in response to membrane thinning [13–15]. The study of the latter systems was prompted by recent observations that transmembranar behavior might be somewhat unstable under Martini 3 [16, 17].

In this work we compare the Martini 2 and 3 behavior of the examined systems to experimental data and observations with finer models to identify aspects in which Martini 3 can be improved. This is relevant as Martini 3 proteins and lipids have not yet been parameterized to fully take advantage of the possibilities of the new model, and our findings can then be considered in upcoming developments.

2. Methods

2.1. Topology generation

All the peptides used in this work are alpha-helical. Their Martini 2.2 and Martini 3 CG topologies were created from atomistic structures (see Table 1 for sequences) using the martini2 [18] script. For the coiled coil systems, the GCN4-leucine zipper mutant N16K parallel homodimer [11] (PDB id: 1ZIK) and the IAAL-E3/K3 synthetic parallel heterodimer [10] (PDB id: 1U0I, model 1/20) were used as starting points; Martini 2 and 3 topologies were generated using angle/torsion restraints to enforce helicity, and for Martini 3 additional topologies were created with elastic restraint networks, using martini2’s default parameters. For the TM WALPs, ideal α-helical structures were created from their sequence using the Avogadro software [19] and the Martini 2/3 topologies constructed using only angle/torsion secondary structure restraints. WALPs are capped on both termini, with N-terminal acetylation and C-terminal n-methyl amidation [13] (Table 1); since neither capping group has been explicitly parameterized for Martini,
we modeled them by simply assigning a non-charged, random-coil, backbone particle type to the WALP’s terminal beads (type P5 in Martini 2 and P2 in Martini 3).

Table 1: Amino acid sequences of the simulated peptides

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>WALP16α</td>
<td>GWW(LA)₃WWA</td>
</tr>
<tr>
<td>WALP19α</td>
<td>GWW(LA)₃LWWA</td>
</tr>
<tr>
<td>WALP23α</td>
<td>GWW(LA)₁₀LWWA</td>
</tr>
<tr>
<td>WALP27α</td>
<td>GWW(LA)₁₀LWWA</td>
</tr>
<tr>
<td>IZIK</td>
<td>RMKQLEDKVEELLSSKKYHLENEVARLKKLV</td>
</tr>
<tr>
<td>1U0I-A</td>
<td>EIAALEKEIAALEKEIAALEK</td>
</tr>
<tr>
<td>1U0I-B</td>
<td>KIALKEKIAALKEIKAALKE</td>
</tr>
</tbody>
</table>

a WALK peptides are termini-capped with N-terminal acetylations and C-terminal α-methyl amidations [13].

2.2. System setup

For the coiled-coil peptides the two monomers comprising each dimer were separated by 4 nm and oriented randomly. Dimers were then solvated in a dodecahedral box of vector length 11.4 nm, with 0.15 M added NaCl, plus neutralizing Cl⁻ counterions for the IZIK system.

The WALP peptides were aligned vertically in order to start the simulation in ideal TM configurations. The peptides were then inserted, using the insane.py script [20], into a 10 × 10 nm² square membrane of DMPC phospholipids (dimyr- toylphosphatidylcholine), with 9 nm simulation box height. Water was added to fill the box, together with an NaCl ionic strength of 0.15 M. In the Martini 2 systems, 10% of water beads were changed to the ‘antifreeze’ type to prevent freezing, as recommended [4].

2.3. Simulation conditions

Simulations were carried out in triplicate, for a minimum 23 µs per replicate, for a total simulation time beyond 2.3 ms (some replicates were simulated significantly longer than others; see Supplementary Table S1). The GROMACS 2020 simulation package [21] was used with a Verlet neighborlist scheme, at a 20 fs time-step. Typical Martini settings were used for non-bonded interactions [22]; Lennard-Jones and Coulombic potentials were cut-off at 1.1 nm, with Coulombic interactions further treated under a reaction-field scheme with a dielectric constant of 15. Temperature was kept at 300 K by a V-rescale thermostat [23] with a 1 ps coupling time. A Parrinello–Rahman barostat [24] with 12 ps coupling time was used to maintain pressure at 1 bar; pressure was coupled isotropically and semi-isotropically (independently in the xy and z directions) for the coiled-coil systems and WALP/membrane systems, respectively. After system assembly and energy-minimization, but prior to production runs, a 2 ns pressure/temperature equilibration was run for all systems, using the more robust Berendsen barostat [25] with a relaxation time of 3 ps.

2.4. System manipulation, analysis and visualization

Peptide alignment and trajectory analysis was done primarily using the MDAnalysis [26, 27] and NumPy [28] Python packages, together with matplotlib [29] for plotting and VMD version 1.9.3 [30] for molecular visualization and rendering. Tilt angles were computed using the HELANAL algorithm [31] as implemented in MDAnalysis. Because the simulated membranes may undergo a small degree of buckling, thicknesses were computed patch-wise, using the binning implemented in the LiPyphilic Python package [32].

3. Results and Discussion

3.1. Coiled coil dimers

The coiled coil is a biologically important structural pattern of helix–helix oligomerization in proteins [12]. The heptad repeat motif is a frequent driver of coiled coils; in it, residues in two or more α-helices interlock in a repeating hypphppp pattern, with hydrophobic residues (h) typically in the first and fourth positions and polar residues (p) in-between (charged residues are often found at the fifth and seventh positions) [12]. The heptad repeat is not a mere polarity matching motif: it leverages the spatial fitting of each helix’s side-chains into inter-residue gaps in the other helices, in what Crick termed a ‘knobs into holes’ pattern [33]. This structural matching further requires helices to become supercoiled, at 3.5 residues per turn, by slightly bending their axes in a left-handed fashion [12].

In our tests, using either Martini 2 or Martini 3 versions, none of the expected coiled coil dimers formed correctly (Figure 1) when compared to the experimentally obtained structures. With Martini 2, dimers did form stably, but incorrectly: the IZIK dimer formed in an antiparallel orientation, contrary to the parallel X-ray structure; the 1U0I dimer formed roughly as frequently in either orientation, with some of the replicate trajectories exhibiting a transition between the two states, but even when correctly parallel the peptides dimerized along a shifted and rotated interface (see the mismatch between the experimental and simulated contacts in Figures 1A and 1B). Martini 2 dimers may have also been overstabilized in that no dissociation was observed in any of the replicates.

In contrast with Martini 2, Martini 3 displayed an interaction propensity between the peptides so low that no discernible predominant dimer structures were observed (Figure 1A). This did not change when the peptides’ secondary structure was restrained by an elastic network instead of angle/dihedral restraints over consecutive residues (see Supplementary Figure S1) — the elastic network alternative is able to maintain the helices supercoiled, though for peptides this short the structural difference is still small. Even when Martini 3 peptides did transiently come into contact, they did so mostly via regions other than the correct dimerization interface.

Our results were somewhat unexpected because of the parameterization focus of Martini in reproducing hydrophobic/hydrophilic preferences. Had this selectivity been correct, the experimental dimerization interfaces should have been recovered, at least in Martini 2. However, the issue of protein
Figure 1: Behavior of coiled coil peptides simulated using the Martini 2 and 3 models. Martini 3 results are shown for topologies without elastic networks (see Supplementary Figure S1 for the very similar behavior with elastic networks). A: Heat maps of contacts between the two monomers (for the 1U0I heterodimer, monomers A and B are in the x and y axes, respectively). Intensity represents the fraction of simulation time across replicates in which residues were within a 6 Å cutoff. Red dots represent experimentally-determined residue contacts (for the 1ZIK X-ray structure, residues within 5 Å; for the 1U0I NMR structure, residues within 4 Å in more than 6 of the 20 deposited minimum-energy models). B: Rendered images of the PDB structures and of representative Martini 2 dimers, with the same reference peptide aligned along a common axis (dashed line) for direct comparison (the parallel Martini 2 dimer was chosen in the 1U0I case); the second peptide is shown darkened, for better distinction. Martini 3 simulations did not indicate any preferred binding mode, and occupancies of the second peptide are represented instead (in pink, for values > 0.1%). Residues are color-coded white for apolar, green for uncharged polar, red for anionic and blue for cationic.

over-aggregation in Martini 2 stems from excessive particle densities and an associated imbalance of effective molecular hydrophobicities/hydrophilicities [2]. Martini 2’s failure to recover the coiled coil polarity matching could be, at least in part, a consequence of such a specific imbalance.

Martini 3 was expected to display weaker peptide–peptide affinities, since one of its development goals was to mitigate the over-aggregation seen in Martini 2. However, to observe almost no dimeric contacts between peptides over an aggregate 80+ µs per system indicates that this type of short proteins now interacts too little. Even transient peptide contacts shunned most of the more hydrophobic correct dimer interface, to instead cluster around polar/charged residues (Figure 1B). This hints at a possible overly hydrophilic character of the peptides. One of Martini 3’s development decisions possibly related to this was the choice to model all backbone beads with the same polar-
ity type (P2), regardless of secondary structure. In Martini 2 the backbone bead types in helical or sheet secondary structures (particle types, N0 and Nda, respectively) were the least polar among backbone types, much less so than the random coil one (P5), and reflected a lower availability of the carbonyl and amide in establishing further hydrogen bonding [5].

In spite of our observations, Martini 3 has been shown to successfully model helix association and the salting-in/salting-out effect of proteins in solution [3]. However, in those validation systems helix association was probed only for TM peptides/proteins; the protein–protein interaction imbalances at play in our tests may be mitigated in a membrane environment, as they already sometimes were in Martini 2 [2, 8, 34]. The salting-in/salting-out Martini 3 tests, however, were performed in solution, and yielded realistic degrees of aggregation for the villin headpiece (VH) and the mutated cellulose-binding domain from *Cellulomonas fimi* (CBD) [3, 35, 36]. Those proteins are only a few kDa larger than the coiled coil dimers in this work but are large enough to have well-defined tertiary structure and globular shape. This makes for larger interaction patches when oligomerizing compared to the almost linear interface between two single α-helices, and may compensate the factors that prevent Martini 3 from properly forming the short coiled coils. Additionally, neither VH nor CBD are fully α-helical/β-sheet structures (helices/sheets account for 70% and 57% of VH and CBD, respectively). If indeed backbone beads are overly polar in Martini 3 α-helices and β-sheets, the relative inaccuracy in overall protein hydrophilicity would then be greater in the 100% helical coiled coil peptides than in VH or CBD. If correct, our explanatory hypotheses for the lack of oligomerization would point to short, fully helical peptides as extreme cases for which Martini 3 selectivity breaks down.

3.2. Transmembrane WALP behavior

TM helices are an important structural aspect of protein–membrane interactions [37]. They are stabilized by such factors as hydrophobic matching (the length matching of helix and membrane hydrophobic spans), or the degree of charge/polarity of the helix’s termini. The fluid nature of interactions between proteins and membranes further means that either component can adapt to non-optimal characteristics [38]; among other changes, excessively long helices – particularly if monomeric or at low degrees of oligomerization – will frequently tilt to better accommodate their hydrophobic span within the membrane core; at the same time, lipids vicinal to the protein may display a thicker bilayer profile, with a concomitant locally thicker core. Conversely, very short helices can induce a local thinning of membranes.

Using Martini 2 and 3 we simulated four α-helices 16, 19, 23, and 27 residues in length, from the WALP (tryptophan–alanine–leucine) series of peptides (termed WALP16, WALP19, WALP23, and WALP27; see Table 1). We chose these peptides because their TM behavior is well established, and their tilt angle as a function of their length has been documented experimentally and with atomistic simulations [13]. This was relevant because we wanted to probe the TM stability of peptides, prompted by recent observations that supposedly transmembranar peptides might, under Martini 3, have weaker affinity for the membrane core [16, 17]. Peptides were simulated embedded in a DMPC membrane, as in the atomistic simulations of ref. 13, above the lipid’s gel-to-fluid transition temperature.

In general, both Martini 2 and Martini 3 were able to recover the length–tilt dependence of WALPs — most visibly for WALP27, which is the longest of the tested peptides (Figure 2). Martini 3 seems to be more sensitive in that WALP23 is also seen to tilt. This is in agreement with reported WALP23 behavior in DMPC [13] and is also an improvement over Martini 2, with which WALP23 does not significantly tilt. The DMPC membrane was also observed to become thinner near the peptides. With Martini 3, this occurred for all peptides, even the longer WALP27, whereas with Martini 2 WALP27 caused a small local thickening instead (see Supplementary Figure S2).

3.2.1. Transmembrane instability

In Martini 3 simulations, WALP16 was observed to consistently be ejected from the membrane core in the µs timescale (Figures 2C and D). This did not occur with Martini 2, and is a behavior that is not supported by experimental results nor by simulations at higher resolutions.

We can preclude a possibly inaccurate thickness of the membrane as the cause for the WALP16 TM instability: due to its coarseness, the Martini 3 DMPC is the same model as the one representing DLPC, and yields a thickness about 2 Å thinner than experimentally reported [39] (3.4 nm vs 3.67 nm; see Supplementary Figure S2). Being somewhat thinner, a Martini DMPC membrane is, if anything, more accommodating of WALP16’s negative size mismatch than if it better matched the experimental thickness (moreover, the mismatch is further reduced by the additional thinning of lipids in the peptide’s vicinity).

Another potential cause for WALP16’s TM-to-adsorbed ejection could be the way we modeled the peptides’ termini. Since neither Martini model has parameters for the WALPs’ N-terminal acetylations or C-terminal α-amidations, we simply considered each terminal bead to be uncharged, and did not add any extra particles. As a consequence, WALPs are marginally shorter than if termini caps had been explicitly modeled, which could have exacerbated the negative size mismatch of WALP16 beyond the point of TM stability. However, we were able to observe the same instability behavior, at a larger timescale, for one of the WALP19 replicates1 (see Figure 2C). If the shortening of termini caps were indeed the destabilizing aspect for WALP16, then WALP19, with 3 extra residues, should have remained transmembranar.

Having ruled out limitations specific to the membrane or WALP models, we are left to conclude that the pattern of Martini 3 protein–membrane affinities is offset in a way that favors

---

1Observation of WALP19 TM instability was serendipitous in that it was first observed, at a shorter timescale, under wrong (isotropic) pressure coupling conditions. This prompted us to monitor WALP19 stability — with correct pressure coupling — into the 10^5 µs scale, and indeed the same behavior was seen as for WALP16.
the adsorbed state vs the TM state, while possibly also lowering the transmembranar-to-adsorbed energy barrier (the point when one of the peptide’s polar ends is dragged across the hydrophobic bilayer core). Note that the same ejection behavior would not result from just the lowering of the energy barrier: were this the case, WALPs would as easily come back to the supposedly more favorable TM state, which is a transition we never observed (Figure 2C).

Our interpretation is in line with recent Martini 3 simulations by Cabezudo and co-workers [17] of membrane self-assembly in the presence of a TM helix dimer. In those simulations, the dimer was effectively excluded from the membrane upon self-assembly. The dimer did remain transmembranar if inserted into pre-assembled membranes, though this is likely to simply reflect a slow TM-to-adsorbed kinetic (in this regard, the simpler WALP systems are advantageous in that ejection events can be observed in the $10^0–10^2$ µs scale).

Cabezudo and co-workers assign the erroneous behavior of Martini 3 to excessively strong peptide–water interactions [17], which parallels the excessive-hydrophilicity hypothesis we put forth above for the lack of aggregation of coiled coil dimers. Indeed, excessive protein–water interactions would selectively stabilize the water-exposed adsorbed state. They would also lower the TM-to-adsorbed energy barrier: as a peptide’s end dives across the membrane core, it progressively allows residues on the opposite end to come into contact with increasing amounts of water; the water interactions of those residues, if too strong, then overly compensate the cost of moving the diving end’s polar moieties through the core. Such a lowering of this energy barrier is likely the reason why we were able to observe TM peptide ejection in the µs scale.

3.3. Mitigation strategies

The correction of the Martini 3 protein–protein and protein–lipid interaction behavior we describe here requires a deeper characterization of the problem(s) at hand and, likely, some degree of re-parameterization of these types of molecules. In the meantime, some mitigating approaches can be employed to enable a more representative simulation of these kinds of systems.

3.3.1. Interaction scaling

By using a scaling factor to adjust cross interactions between molecules, or between them and water, one can centrally re-balance the affinities of an entire system. This strategy is not new, having been proposed to address the excessive protein aggregation under Martini 2 — either by decreasing protein–protein interactions [7, 8] or by increasing protein–water ones [9].
Cabezudo and co-workers have proposed the scaling-down of protein–water interactions to stabilize Martini 3 TM peptide states [17]. The fact that they successfully employed this strategy reinforces that, indeed, a hydrophobic/hydrophilicity imbalance is at play. The same strategy can, in principle, be applied to the coiled coil dimer systems: weakening protein–water interactions would certainly drive a stronger dimerization. It is unclear, however, whether a blanket scaling would be able to correct the apparent rejection of hydrophobic contacts seen in Figure 1B.

Interaction scaling has the downside of impacting all adjusted interactions, and not only the ones responsible for the model’s misbehavior. When defining scaled interactions, one also typically does not repeat the extensive testing that went into the force field’s parameterization — the point being precisely to avoid such an in-depth redo. Because of these aspects, solutions involving interaction scaling have questionable transferability properties, and are better seen as temporary stopgap measures for the specific systems they were developed for.

3.3.2. System restraining

A second approach to drive simulations towards expected behavior is to use some sort of potential that restrains the phase space sampled by the simulations. In a general sense, this is already the case of angle and dihedral bonded interactions in Martini, that account for the fact that the rest of CG interactions represent poorly intramolecular dynamics on their own.

A typical restraining strategy applied to Martini protein systems is the so-called elastic network [40], where several potentials restrain a set of distances between backbone particles to predefined values. Depending on their reach, elastic networks can restrain from secondary up to quaternary protein structures, and are therefore an adequate solution to simulating coiled coils as dimers.

With unstable TM peptides, restraints preventing the termini from diving into the bilayer core have also been successfully used [16]. This also requires restraining the membrane from buckling or from otherwise shifting significant, else it may happen that the membrane becomes the one to leave the peptide.

The use of restraints has the advantage that the transferability properties of the model are generally not affected. This is particularly true if the system is restrained to a state that was already a local energy minimum and therefore only requires sporadic biasing (such as a TM peptide). The downsides of this approach relate to the loss of freedom in the restrained degrees of freedom: dynamic behavior involving the restrained dimensions becomes an input of the system, and no longer something that can be explicitly modeled or inferred. For instance, with elastic-networked coiled coils one would be unable to comment on monomer affinity or on their dynamics of pairing.

4. Conclusion

With our work we characterize important aspects of Martini 3 protein–protein and protein–membrane interactions and identify apparent shortcomings of the force field. By drawing attention to these points, and discussing several mitigation strategies, we aim to advise others studying related systems and hopefully limit the waste of researcher time and processor cycles on unproductive simulations.

Finally, given its young age, Martini 3 can be expected to still undergo several improvements. We hope our findings are a step towards such future developments.

5. CRediT authorship contribution statement

J. Karl Spinti: Data curation, Writing - Original draft preparation, Formal analysis, Visualization. Fernando Neiva Nunes: Data curation, Writing - Original draft preparation, Formal analysis, Visualization. Manuel N. Melo: Conceptualization, Supervision, Writing - Original draft preparation, Writing - Review & Editing, Methodology, Visualization.

6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7. Data availability

Files used for simulating the systems, as well as contact and tilt-angle data in numerical form, are available for download from the Melo lab’s website, at https://www.itqb.unl.pt/labs/multiscale-modeling.

8. Acknowledgements

We thank T. Cordeiro for bringing to our attention the coiled coil system that motivated part of this study. J.K.S. acknowledges an internship sponsored by Fundação Luso-Americana para o Desenvolvimento through its Study in Portugal Network. M.N.M. thanks Fundação para a Ciência e a Tecnologia for fellowship CEECIND/04124/2017, and for funding project MOSTMICRO-ITQB with references UIDB/04612/2020 and UIDP/04612/2020.

References
