On the Automatic Optimization of Lipid Models in the Martini Force Field using SwarmCG

Charly Empereur-mot,[a,*] Kasper B. Pedersen,[b] Riccardo Capelli,[c] Martina Crippa,[d] Cristina Caruso,[d] Mattia Perrone,[d] Paulo C.T. Souza,[e] Siewert J. Marrink[f] & Giovanni M. Pavan[a,d,*]

[a] *Department of Innovative Technologies, University of Applied Sciences and Arts of Southern Switzerland, Polo Universitario Lugano, Campus Est, Via la Santa 1, 6962 Lugano-Viganello, Switzerland*

[b] *Department of Chemistry, Aarhus University, Langelandsgade 140, Aarhus C, Denmark*

[c] *Department of Biosciences, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy*

[d] *Politecnico di Torino, Department of Applied Science and Technology, Corso Duca degli Abruzzi 24, 10129 Torino, Italy*

[e] *Molecular Microbiology and Structural Biochemistry (MMSB, UMR 5086), CNRS & University of Lyon, 7 Passage du Vercors, 69007 Lyon, France*

[f] *Molecular Dynamics, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands*

[*] *Corresponding authors:* charly.empereur-mot@supsi.ch, giovanni.pavan@polito.it

ABSTRACT

After two decades of continued development of the Martini coarse-grained force field (CG FF), further refining the already rather accurate Martini lipid models has become a demanding task that could benefit from integrative data-driven methods. Automatic approaches are increasingly used in the development of accurate molecular models, but they typically make use of specifically-designed interaction potentials that transfer poorly to molecular systems or conditions different than those used for model calibration. As a proof of concept here we employ SwarmCG, an automatic multi-objective optimization approach facilitating the development of lipid force fields, to refine specifically the bonded interaction parameters in building blocks of lipid models within the framework of the general Martini CG FF. As targets of the optimization procedure, we employ both experimental observables (top-down references: area per lipid & bilayer thickness) and all-atom molecular dynamics simulations (bottom-up reference), respectively informing on the supra-molecular structure of the lipid bilayer systems and on their sub-molecular dynamics. In our training sets we simulate at different temperatures in the liquid and gel phases up to 11 homogeneous lamellar bilayers, composed of phosphatidylcholine lipids spanning various tail lengths and degrees of (un)saturation. We explore different CG representations of the molecules and evaluate improvements a posteriori using additional simulation temperatures and a portion of the phase diagram of a DOPC/DPPC mixture. Successfully optimizing up to ~80 model parameters within still limited computational budgets, we show that this protocol allows to obtain improved transferable Martini lipid models. In particular, the results of this study demonstrate how a fine tuning of the representation and parameters of the models may improve their accuracy, and how automatic approaches such as SwarmCG may be very useful to this end.

I. INTRODUCTION

Molecular dynamics (MD) has become a cornerstone tool in the study of complex molecular systems, providing high-resolution insights often inaccessible via experimental techniques. Coarse-grained (CG) molecular modeling, in which atoms are bundled together into supra-atomic particles, extends the space and time scales accessible via MD simulations and is increasingly employed to characterize systems of interest in structural biology, drug discovery, biophysics and nanomaterials design[1,2]. Martini[3–5] is a popular CG modeling scheme, which provides pre-parametrized molecular fragments (beads) for creating molecular models in an additive fashion. Non-bonded interactions between CG beads are described via simple spherical Lennard-Jones (LJ) and Coulomb potentials, parametrized according to the miscibility and partitioning of their associated molecular fragment between different solvent environments (*top-down* route), while bonded interactions are usually calibrated based on equilibrium simulations of higher-resolution molecular models (*bottom-up* route). The resulting simplification of the molecular systems enables a speed up of 2 to 3 orders of magnitude with respect to equivalent all-atom (AA) modeling[2].

The recent re-parametrization of the Martini force field (FF, version 3.0.0[4]) improved the overall balance of *non-bonded* interactions between beads, as well as the accuracy of the scheme in predicting molecular packing in MD simulations. Notably, enhanced CG representation of molecular volumes can be obtained via the use of higher-resolution particles (small and tiny bead sizes). Particular attention was paid to the description of aliphatic and aromatic ring-like structures, which are ubiquitous in small molecules (*e.g.* solvents, drugs) and building blocks constituting macromolecules (*e.g.* proteins, synthetic polymers). Such improvements enable the modeling of increasingly complex systems comprising multiple classes of molecules, such as solvent mixtures, small molecules, polymers, lipid membranes, proteins and protein-ligand systems; all within the framework of a general force field[6–17].

In this paper, we focus on lipid models which remain yet to be updated to fully take advantage of the new interaction matrix available in Martini 3.0.0[4]. Notably, their CG representations have remained mostly conserved since the inception of the FF[3,18] and do not allow to differentiate between some of the lipids. This is the case, *e.g.*, of phosphatidylcholine (PC) lipids DLPC and DMPC, respectively including 12 and 14 carbons per "tail" and currently represented by identical CG models (Fig. 1c), while their phosphate-to-phosphate bilayer thickness (D_{HH}) differ by ~15% at room temperature[19,20]. The exclusive usage of big beads for modeling the fatty acids, which are designed to represent 4 heavy atoms and their associated hydrogens, does not provide enough resolution to differentiate between the two lipids. Further refining the CG representation of the models, in principle, will allow to further enhance thermodynamic properties of lipids and lipid mixtures in Martini simulations[4,21]. New experimental measurements have also become available recently for lipids containing polyunsaturated fatty acids (PUFA) and should be considered for guiding the refinement of their respective models (*e.g.* D_{HH} thickness for SDPC and PDPC)[22].

Then, among the wide variety of thermotropic phases exhibited by saturated lipid membranes, the tilted gel ($L_{\beta'}$) and ripple ($P_{\beta'}$) phases are not accurately described using current Martini models, as bilayers preferentially adopt exclusively the straight gel phase (L_{β}) at low temperatures[4,23]. The further refinement of these models could focus on enabling better stabilization of the tilted gel phase ($L_{\beta'}$), as well as triggering phase transitions at their relevant temperatures in simulations[4,24]. Lastly, although only sparse data is available, experimental phase diagrams of lipid mixtures constitute one of the most information-rich sources usable in the calibration or validation of a FF. Maximizing the fidelity of CG simulations to phase diagrams of lipid mixtures, however, remains a challenging endeavor both in terms of FF calibration effort and computational effort (*i.e.* computational time, elaborate simulation setups)[4,24]. After two decades of continued Martini development, further refining the lipid models has become a demanding task that could benefit from automatized procedures and machine learning[2].

Here we employ *SwarmCG*[25,26], an automatic multi-objective optimization approach that facilitates the development of transferable lipid FFs, to evaluate and compare the potential of two putative refined CG representations of the Martini lipids. Using simultaneously in the training sets 8 PC lipids simulated at different temperatures and spanning different phase states, we optimize the *bonded* parameters in building blocks of the lipid models while *non-bonded* parameters remain constant (set to Martini 3.0.0[4]). Applying *SwarmCG*[25,26] here guarantees that an optimal set of *bonded* parameters has been found for each resolution compared, according to a given set of molecular modeling constraints (*i.e.* composition of the training set, *non-bonded* parameters and a limited number of other simulation parameters). This eliminates uncertainties related to parameters tuning, which in turn allows to gain insights on the relative ability of each CG representation to further enhance the thermodynamic properties of the lipid models in the Martini framework. The two putative CG representations are compared *a posteriori* using a range of temperatures and by simulating a portion of the phase diagram of a DOPC/DPPC mixture known to simultaneously exhibit two phase states experimentally[24,27].

II. METHODOLOGY

A. Automated optimization of CG lipid models

The optimization protocol proposed in *SwarmCG*[25,26] allows to obtain CG FF parameters for lipid models by simultaneously exploiting experimental data (*top-down*: area per lipid & D_{HH} thickness) and AA MD simulations (*bottom-up:* bond & angle distributions), respectively informing on the supra-molecular structure of the systems and on their sub-molecular dynamics (Fig. 1a). CG models are tested iteratively in a set of simulations of small patches of homogenous lamellar bilayers, used to measure the discrepancies observed between the simulated and reference data (128 lipids and 200 ns of equilibrium CG MD simulation each). The software allows to modulate selected parameters of the CG FF, that are iteratively optimized within pre-defined boundaries using FST-PSO[28] (Fuzzy Self-Tuning Particle Swarm Optimization, one of the most efficient PSO variant to date[29]) for minimizing a loss function designed to improve FF accuracy.

The quality and completeness of the information embedded in the training set directly conditions the accuracy of the output FF, as well as its ability to generalize to different types of lipids and to different thermodynamic conditions (than those included in the training set). In this study, because the *bonded* parameters are defined as general building blocks and redundant across lipid types (Fig. 1e), while at the same time the space of their potential solutions is restrained by other simulation parameters remaining constant (notably *non-bonded* parameters set to Martini 3.0.0[4]), using a rich training set can be expected to output a consistent CG lipid FF that generalizes well to other types of lipids and to other observables than those used as objectives during optimization[26]. The CG representations of the lipids, however, might still limit the extent to which reference data can be fitted, depending on the degrees of freedom preserved by the molecular models (*i.e.* depending on the choice of beads describing *non-bonded* interactions, their reference positioning, the topology of the CG models and the potentials used to describe *bonded* interactions).

In the following experiments, we use in the training sets 8 PC lipids spanning tails of different lengths and different degrees of (un)saturation, for which experimental measurements are available for pure composition lamellar bilayers in the liquid and gel phases (area per lipid and D_{HH} bilayer thickness)[19,20,22,30]. Among the corresponding models in Martini, the following pairs are currently represented identically: DLPC & DMPC (12:0-12:0 & 14:0-14:0), DPPC & DSPC (16:0-16:0 & 18:0-18:0), POPC & SOPC (16:0-18:1 & 18:0-18:1) and SDPC & PDPC (16:0-22:6 & 18:0-22:6) (Fig. 1c). This is also the case for their respective head-type variants: DLPE & DMPE, DLPS & DMPS, *etc.* To resolve this issue we explore two different CG representations of the lipids that allow to uniquely map each molecule, experimenting with the use of multiple bead sizes to preserve degrees of freedom in the models and allowing *SwarmCG*[25,26] to precisely accommodate the reference data.

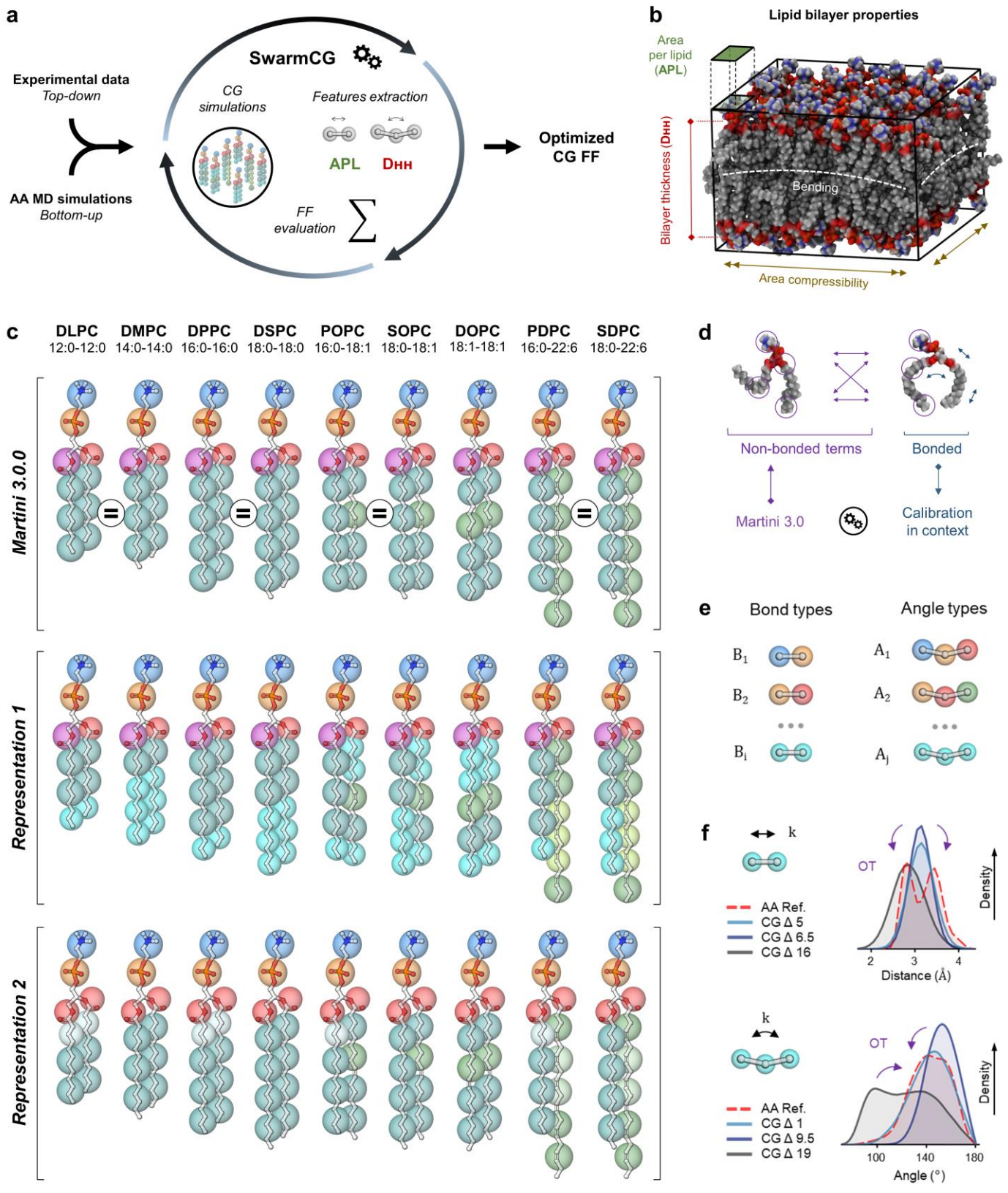

Figure 1. Overview of the protocol followed for obtaining *bonded* parameters via *SwarmCG* for different CG models of lipids within the framework of Martini 3.0.0. (a) *SwarmCG* simultaneously relies on *bottom-up* and *top-down* references to iteratively optimize model parameters using higher-resolution AA MD simulations and experimental data. (b) Illustration of lipid bilayer properties showing notably the APL and D_{HH} used for calculating the *top-down* component of the loss function. (c) Overview of the CG representations of interest in this study with CG beads mapping shown over AA structures, using beads Q1 (dark blue), Q5 (orange), SN4a (red), N4a (purple), C1 (blue), SC2 (cyan), SC1 (white), C4h (olive), C5h (light green) and SC4h (bright yellow/green). (d) Principle of the parameters calibration in this study: *bonded* parameters of the models are calibrated in the context of *non-bonded* interaction terms set to Martini 3.0.0, iterating CG MD simulations of bilayers composed of different types of lipids.

(e) CG bonds and angles are defined as building blocks and classified according to the CG beads they involve, which defines the type of a specific bond/angle as well as the parameters employed. (f) Principle of the OT-B metrics used for exploiting structure-based information from AA reference simulations.

B. Loss function

The loss function, to be minimized, reformulates our many-objective optimization problem into a single-objective one by aggregating the distances from both *top-down* and *bottom-up* objectives into a global "FF accuracy score" (loss global in Fig. 2b and Fig. 4b). Fitting the APL and DHH experimental data is formulated as the two primary objectives, having equal importance (*top-down* global features), while reproducing within the CG models the distributions of bond and angle values calculated from AA MD simulations constitutes a secondary objective (*bottom-up* local features). To this end, we apply a soft penalty which role is to ensure large deviations from the reference *bottom-up* data are forbidden and allows to effectively distribute in between the *bonded* building blocks the residual error inherent to the coarse-graining process (*i.e.* inherent to the reduction of the number of degrees of freedom in between AA and CG models). The loss function is defined as

$$loss = \sqrt{\left(\Delta APL_{global}{}^2 + \Delta D\mathrm{HH}_{global}{}^2 + OT\text{-}B_{global}{}^2\right)/3}, \quad (1)$$

where ΔAPL_{global} and $\Delta D\mathrm{HH}_{global}$ are the aggregated percentage deviations from experimental data, calculated across the training set as

$$\Delta APL_{global} = \sqrt{\sum_i^n \left(w_1 + max(0, |\Delta\%APL_i| - \varepsilon)\right)^2 / n}, \quad (2)$$

and

$$\Delta D\mathrm{HH}_{global} = \sqrt{\sum_i^n \left(w_1 + max(0, |\Delta\%D\mathrm{HH}_i| - \varepsilon)\right)^2 / n}, \quad (3)$$

where $\Delta\%APL_i$ and $\Delta\%D\mathrm{HH}_i$ are the percentage deviations observed on average within the i^{th} simulation of a given pair of lipid type and temperature included in the training set, ε represents the tolerated measurements error in $\Delta\%APL_i$ and $\Delta\%D\mathrm{HH}_i$ (set to 1.5), w_1 is a weight used to prioritize fitting the *top-down* objectives over the *bottom-up* one (set to 10), meaning the protocol is allowed to discard structure-based information for better fitting experimental measurements, and n is the number of pairs of lipid type and temperature included in the training set.

For calculating the *bottom-up* component of our loss function ($OT\text{-}B_{global}$ in Eq. 1), as preliminary steps we obtain well-sampled AA MD trajectories for each bilayer included in the training set (128 lipids and 1 μs of equilibrium AA MD simulation each). We then project (map) these trajectories according to our CG representations of interest (Fig. 1c) and compute all AA-mapped bond and angle distributions to be used as reference during optimization. The bond and angle types are defined respectively as all possible connected pairs and triplets of beads, mapping atomic connectivity into beads connectivity according to our CG

representations (Fig. 1c,e). We evaluate the mismatch between corresponding CG *vs.* AA-mapped bond and angle distributions using the Wasserstein distance[31,32] (a.k.a. Earth Movers' Distance, EMD), a metric based on optimal transport[33] (OT, Fig. 1f) having several useful properties: (i) multimodal distributions are well handled; (ii) distances are robust to noise; (iii) distances are quantified in interpretable units (*e.g.*, Å, degrees); and (iv) their computations are inexpensive (via PyEMD[31,34]). This metrics (hereafter referred to as "OT-B metrics") has been already proven well-suited for parametrizing the *bonded* terms of CG models of complex and flexible molecules in a previous version of *SwarmCG*[25]. The *bottom-up* component of the loss informs on how closely a putative set of FF parameters allows to match the AA description of the molecular systems included in the training set, and is calculated as

$$OT\text{-}B_{global} = \sqrt{\left(\sum_b^B \sqrt{\frac{\sum_i^n (w_2 \times OT\text{-}B_{b,i})^2}{n}} + \sum_a^A \sqrt{\frac{\sum_j^m OT\text{-}B_{a,j}{}^2}{m}}\right) \Big/ (B+A)}, \quad (4)$$

where $OT\text{-}B_{a,i}$ and $OT\text{-}B_{b,j}$ are the OT-B distances calculated respectively for the i^{th} instance of bond type b and j^{th} instance of angle type a, with n and m respectively the number of instances of bond type b and angle type a in all simulations included in the training set, and B and A respectively the number of bond and angle types appearing in the CG models included in the training set (Fig. 1e). The weight w_2 (set to 50) allows to obtain comparable OT-B values whether the metrics is applied on bonds or angles, and prioritizes adjusting first the distributions of the bonds. Setting w_2 to 50, a distance of 0.4 Å between the distributions of two bonds is considered equivalent to a distance of 20 degrees between the distributions of two angles. Setting w_1 to 10, fitting the experimental measurements is prioritized as long as OT-B distances are below 0.2 Å for bonds and below 10 degrees for angles (on average and for all bond and angle types). The convergence criterion is defined as 10 swarm iterations without improving loss.

III. RESULTS

A. Representation 1: Mixed tail resolution helps improving fidelity in the liquid phase

As a first experiment, we employ *SwarmCG*[25,26] to calibrate the *bonded* parameters in building blocks of PC lipid models using Representation 1 (Fig. 1c), including in the training set lipids spanning tails of various lengths and degrees of unsaturation. The hypothesis motivating this CG representation is that mapping the lipid tails as precisely as possible, using regular beads to represent exactly 4 heavy atoms and small beads to represent exactly 3 heavy atoms, may allow an enhanced description of their flexibility and dynamics (*n.b.* exclusively regular beads were used to represent lipid tails thus far and Martini 3.0.0[4] offers well-calibrated smaller beads). We optimize model parameters by iteratively simulating 9 different patches of lamellar bilayers, having each of the 8 lipid types simulated in the liquid phase and only the DPPC bilayer simulated also in the tilted gel phase ($L_{\beta'}$) at 293K. We calibrate equilibrium values and force constants for all bonds and all angles defined in the lipid models, totaling 77 parameters across 16 bond types and 27 angle types

(Fig. S1, equilibrium values remain set to 180 degrees for 9 angle types). The swarm of particles is initialized randomly within the 77-dimensional search space, except for the first particle that is provided with knowledge from AA-mapped MD simulations (initialized using average equilibrium values computed for all bond and angle types) and from previous Martini lipid models[4] (initialized within relevant ranges of force constants).

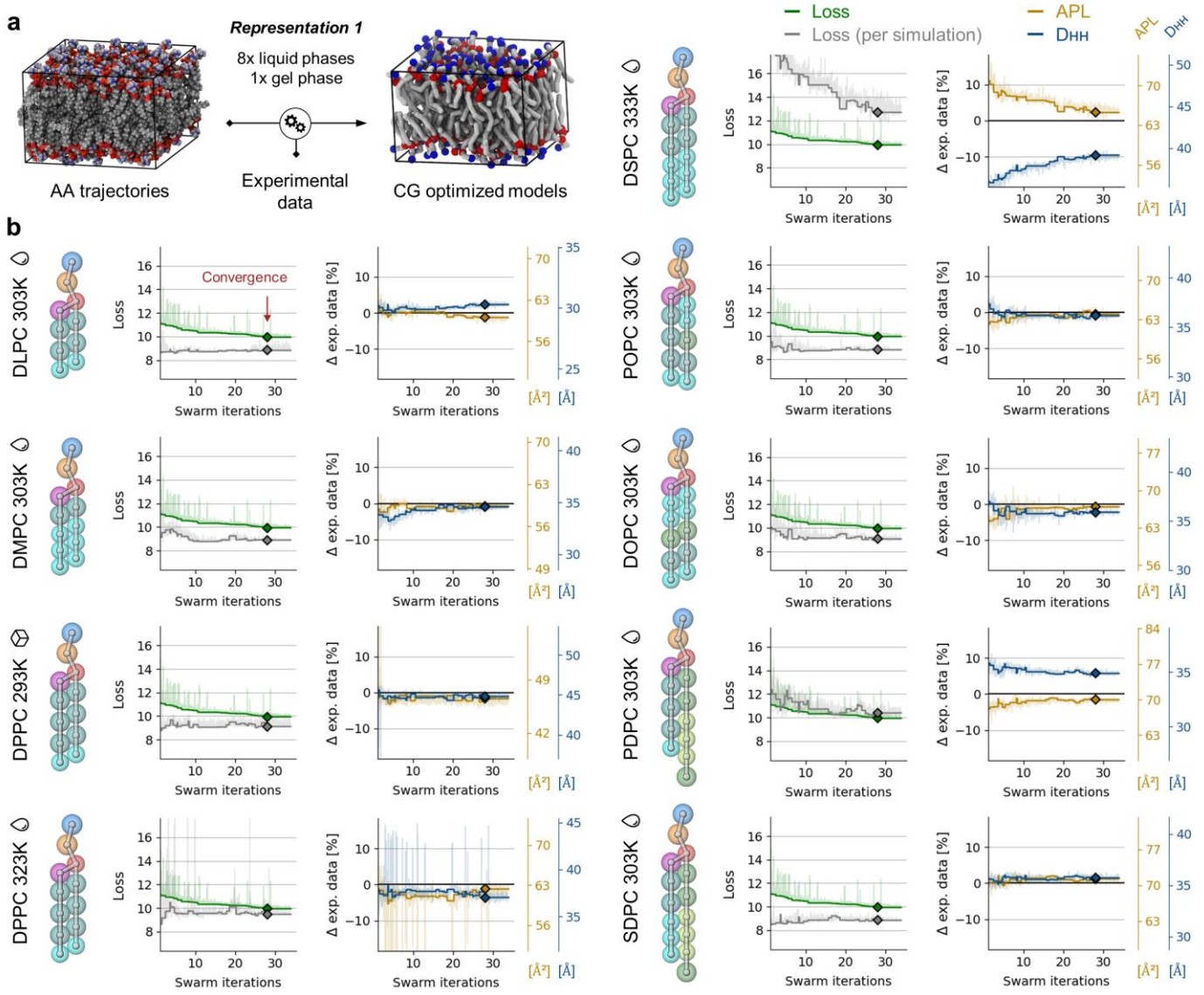

Figure 2. Multi-objective optimization of the *bonded* parameters of the FF for PC lipid models built in the framework of Martini 3.0.0 using Representation 1 and in the training set bilayers of 8 different lipid types simulated at 9 temperatures (DLPC 303K, DMPC 303K, DPPC 293K & 323K, DSPC 333K, POPC 303K, DOPC 303K, PDPC 303K and SDPC 303K). (a) Illustration summarizing the workflow. (b) Left panels: loss global (green) and loss per bilayer simulation (grey) in the training set. Right panels: APL (yellow) and DHH (blue) for each bilayer simulation in the training set. The horizontal black lines set at 0 identify the target experimental APL and DHH values. Solid curves are values corresponding to the best global loss at any point during optimization. Shaded lines show raw data. Diamonds represent values at convergence, obtained with the optimized *bonded* parameters.

The steady decrease of the loss (Fig. 2b: green curve) indicates the *bonded* parameters of the models are adapted successfully and allow approaching the objectives, until the optimization converges at swarm iteration 29. At convergence, the models overall correctly reproduce the APL and DHH experimental

measurements defined as target (Fig. 2b: yellow and blue curves converging towards reference black line, set to 0) and only the DHH values for DSPC and PDPC in the liquid phase remain meaningfully improvable (*n.b.* here we calculate the phosphate-to-phosphate bilayer thickness to approximate the DHH). This is also visible when calculating the loss separately for each simulation (Fig. 2b: grey curves), which additionally indicates that fitting the objectives is harder for DSPC at 333K than it is for PDPC at 303K (Fig. 2b: grey curves remain over green curves) in the context of the modeling choices of this experiment (CG representation, composition of the training set, *etc.*). The OT-B distances otherwise are minimized effectively, indicating that the structural features present in the reference AA-mapped MD trajectories are overall well reproduced in the CG descriptions of the systems (Figs. S2-4). In terms of computational time, here the refinement of 77 *bonded* parameters of the lipid models using 9 informative simulations required 15 days (wall-clock time) to reach convergence (39 swarm iterations) using 27 particles in the swarm and using 64 CPU cores (each CG simulation running on a single CPU core, scaling horizontally by parallelizing the swarm of particles on an inexpensive CPU machine).

To further estimate the balance of *bonded* and *non-bonded* interactions in these models, we evaluate *a posteriori* their ability to describe phase separation in DOPC/DPPC mixtures known to simultaneously exhibit two phase states at 298K[24,27] (this phase separation is not correctly described using Martini lipid models version 3.0.0). Simulating a bilayer composed of 1152 randomly dispersed lipids at 10/90% mass DOPC/DPPC and starting from a configuration in the liquid phase, the nucleation of the gel phase is observed after ~100 ns of equilibration and the equilibrated system exhibits liquid/gel phase separation after ~1 μs (Fig. 3a). To characterize the phase separation we employ the LENS descriptor[35], which allows to evaluate locally the changes in the environment of a molecule across trajectory frames and here enables classifying lipids into phase states. The proportion of lipids in the gel and liquid phases is stable in an additional 10 μs of production simulation, during which the two phases constantly exchange lipids (Fig. 3c: stars, Supplementary Movie 1). Re-iterating this experiment with 20/80% mass DOPC/DPPC in the system, we did not observe gel phase nucleation after 1 μs of equilibration. Lowering the simulation temperature to 293K, stable phase separation could be observed also at 20/80% mass DOPC/DPPC after ~1 μs of equilibration (Fig. 3b), again with a constant exchange of lipids in between the two phases (Fig. 3c: squares).

Altogether these results indicate that an optimal solution was found for parametrizing the *bonded* terms of Martini lipid models using the alternative Representation 1, according to the objectives defined. Structural properties appear improved for the selected lipids in the liquid phase and in the gel phase specifically for DPPC, the *bonded* parameters having been calibrated in order to obtain an optimal compromise specifically to this end. However, because Representation 1 makes use of repeated small and big beads for modeling saturated tails (Fig. 1c), the lateral packing of the tails cannot be correctly described in the gel phase notably for DMPC and DSPC. In particular, when tested *a posteriori* the transition in between liquid and gel phases is not observed for DMPC and the transition temperature is not respected for DSPC (Fig. 5b). Repeating this

parametrization experiment and including in the training set also simulations of DMPC and DSPC bilayers in the gel phase (273K and 308K, respectively) does not allow to further improve the models. Instead, this additional experiment shows that no satisfying solution can be found for parametrizing towards these objectives the *bonded* parameters of lipid models using the philosophy of Representation 1 (Fig. S5).

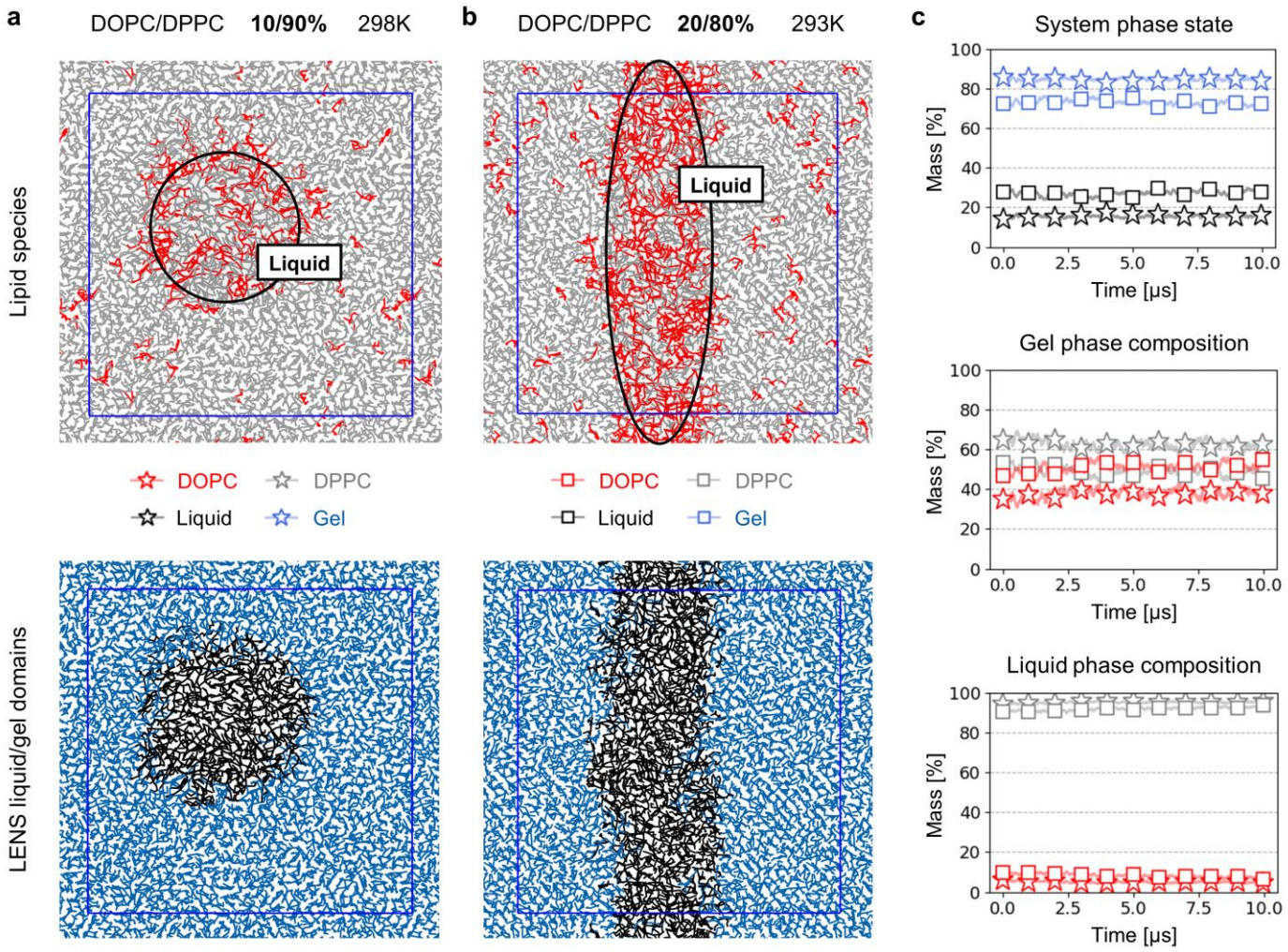

Figure 3. Characterization of the phase separation in DOPC/DPPC mixtures with lipid models using Representation 1. (a) Orthogonal view of a bilayer composed of 1152 lipids at 10/90% mass DOPC/DPPC. Top: Colored according to phase state using LENS (blue: gel phase, black: liquid phase). Bottom: Colored according to lipid type (red: DOPC, grey: DPPC). (b) Orthogonal view of a bilayer composed of 1152 lipids at 20/80% mass DOPC/DPPC. Top: Colored according to phase state using LENS (blue: gel phase, black: liquid phase). Bottom: Colored according to lipid type (red: DOPC, grey: DPPC). (c) Left: Mass percentage of the system in the gel (blue) and liquid (black) phase across simulations at 10% (stars) and 20% (squares) mass DOPC. Middle: Mass percentage of DOPC (red) and DPPC (grey) found in the gel phase across simulations at 10% (stars) and 20% (squares) mass DOPC. Right: Mass percentage of DOPC (red) and DPPC (grey) found in the liquid phase across simulations at 10% (stars) and 20% (squares) mass DOPC.

B. Representation 2: More homogeneous tail resolution balances improvement across phases

We then adapt our protocol and perform an automated search of an optimal set of *bonded* parameters for calibrating the building blocks of lipid models using Representation 2 (Fig. 1c). The hypothesis motivating

this alternative representation is that while attempting to better preserve degrees of freedom during the coarse-graining process, using exclusively regular beads in the bulk of all saturated tails remains mandatory for preserving the thermodynamic properties of the lipids in both the liquid and gel phases (*n.b.* tails are differentiated by using next to ester groups either a small bead with 3-to-1 mapping or a regular bead with 5-to-1 mapping). We include in the training set the same 8 PC lipids as in the previous experiment, for which we simulate patches of lamellar bilayers in the liquid phase at the same temperatures as previously, but this time we directly include bilayer simulations in the tilted gel phase ($L_{\beta'}$) for DMPC, DPPC and DSPC (273K, 293K and 308K, respectively). This provides more information for guiding the optimization and introduces additional constraints enforcing the transferability of the *bonded* building blocks across phase states. We calibrate equilibrium values and force constants for all bonds and all angles defined in the lipid models, totaling 48 parameters across 13 bond types and 12 angle types (Fig. S6, equilibrium values remain set to 180 degrees for 2 angle types). The swarm of particles is initialized randomly within the 48-dimensional search space, except for the first particle that is provided with knowledge from AA-mapped MD simulations (initialized using average equilibrium values computed for all bond and angle types) and from previous Martini lipid models (initialized within relevant ranges of force constants known to be adequate).

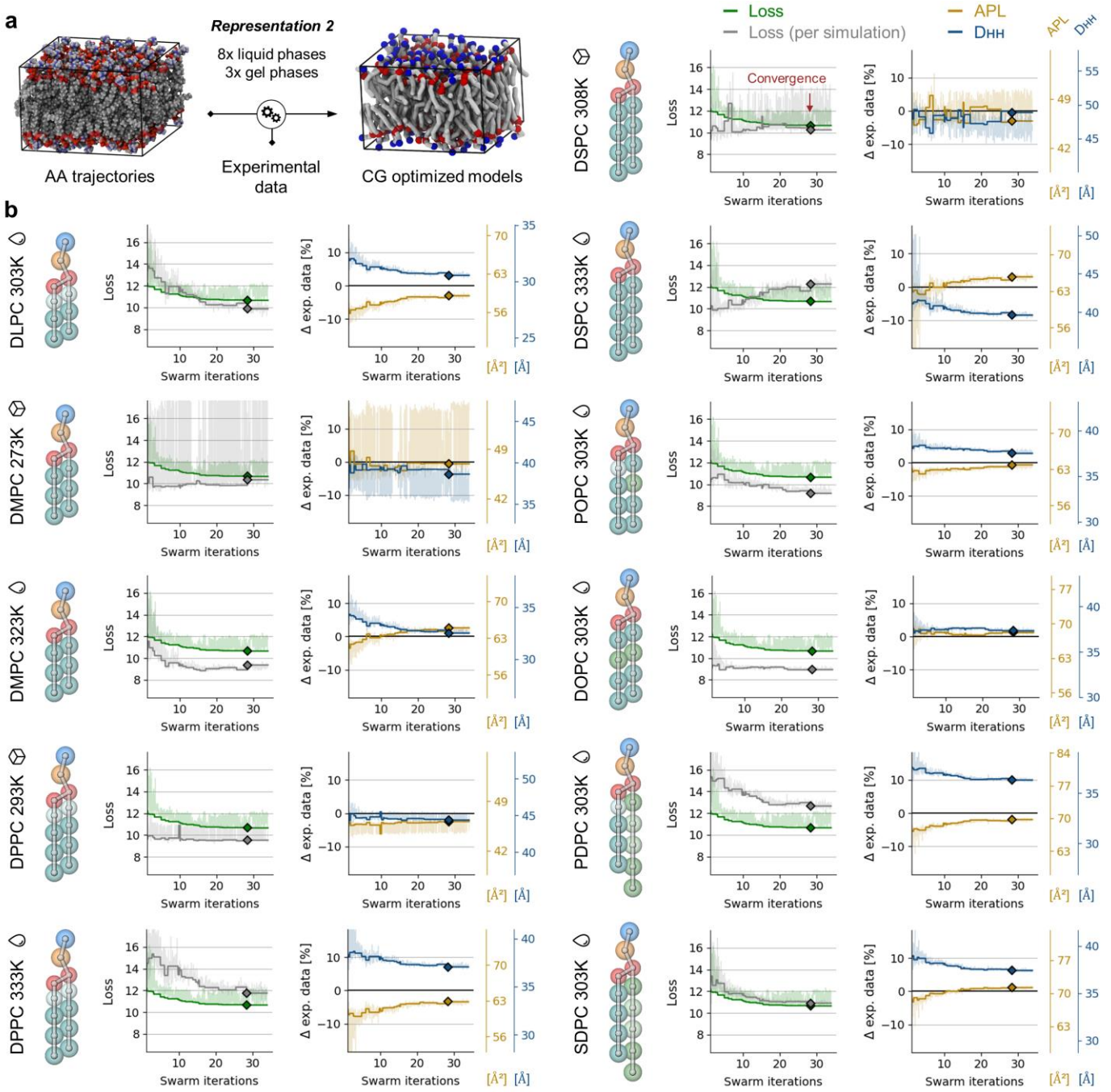

Figure 4. Multi-objective optimization of the *bonded* parameters of the FF for PC lipid models built in the framework of Martini 3.0.0 using Representation 2 and in the training set bilayers of 8 different lipid types simulated at 11 temperatures (DLPC 303K, DMPC 273K & 303K, DPPC 293K & 323K, DSPC 308K & 333K, POPC 303K, DOPC 303K, PDPC 303K and SDPC 303K). (a) Illustration summarizing the workflow. (b) Left panels: loss global (green) and loss per bilayer simulation (grey) in the training set. Right panels: APL (yellow) and D_{HH} (blue) for each bilayer simulation in the training set. The horizontal black lines set at 0 identify the target experimental APL and D_{HH} values. Solid curves are values corresponding to the best global loss at any point during optimization. Shaded lines show raw data. Diamonds represent values at convergence, obtained with the optimized *bonded* parameters.

The steady decrease of the loss (Fig. 4b: green curve) indicates the *bonded* parameters of the models are adapted successfully until the optimization converges at swarm iteration 28. The presence of noise in the evaluation of the APL and D_{HH} thickness for DMPC at 273K and DSPC at 308K indicates that the sets of parameters explored during optimization did not systematically trigger the formation of a gel phase (Fig. 4b: light yellow and blue curves). At convergence, the models overall correctly approach the APL and D_{HH}

experimental measurements defined as target (Fig. 4b: yellow and blue curves converging towards reference black line, set to 0), indicating Representation 2 allows to describe the thermodynamic properties of various types of lipids. The optimization problem now includes more and different constraints on the *bonded* parameters, introduced by adding in the training set the simulations of DMPC and DSPC in the tilted gel phase ($L_{\beta'}$), as well as by reducing the number of beads in the lipid tails notably for PUFA-containing lipids (*i.e.* slightly lower resolution than Representation 1). The APL is very well fitted to experimental data for all pairs of lipid type and simulation temperature, both in the liquid and gel phases, while the D_{HH} thicknesses remain slightly less fitted on average. The OT-B distances here also are minimized effectively, overall reproducing within the CG models the structural features present in the reference AA-mapped MD trajectories (Figs. S7-9).

The error on fitting D_{HH} experimental data remains the largest for DPPC and DSPC at 333K and for PDPC and SDPC at 303K (Fig. 4b: blue curves reaching a plateau away from reference black line). These four lipids have in common saturated tails including 16 and 18 carbons (PDPC and DPPC both include 16:0 tails and SDPC and DSPC both include 18:0 tails), indicating compromises had to be made during the optimization of the *bonded* parameters of these tails for matching experimental data set as target in different phase states and for the highly flexible PUFA-containing lipids. In the context of this rich training set, representative of the variety of PC lipids and including transversal experimental data, this result indicates there exists no set of *bonded* parameters allowing to further improve the matching of D_{HH} thicknesses for these lipid models using Representation 2 (in the context also of the other optimization and modeling constraints: ranges defined for the exploration of equilibrium values and force constants, bonds and angles defined in the CG representation, *non-bonded* parameters and other simulation parameters used). Calculating the loss separately for each simulation shows that fitting the objectives was harder particularly for DSPC at 333K and PDPC at 303K (Fig. 4b: grey curves remain over green curves). Repeating the previous simulation protocol of a DOPC/DPPC mixture at 10/90% mass for Representation 2, we could observe gel/liquid phase separation after adapting the temperature to 288K (10K below the experimental reference). In terms of computational time, here the refinement of 48 *bonded* parameters of the lipid models using 11 informative simulations required 13 days (wall-clock time) to reach convergence (38 swarm iterations) using 23 particles in the swarm and using 64 CPU cores (each CG simulation running on a single CPU core, scaling horizontally by parallelizing the swarm of particles on an inexpensive CPU machine).

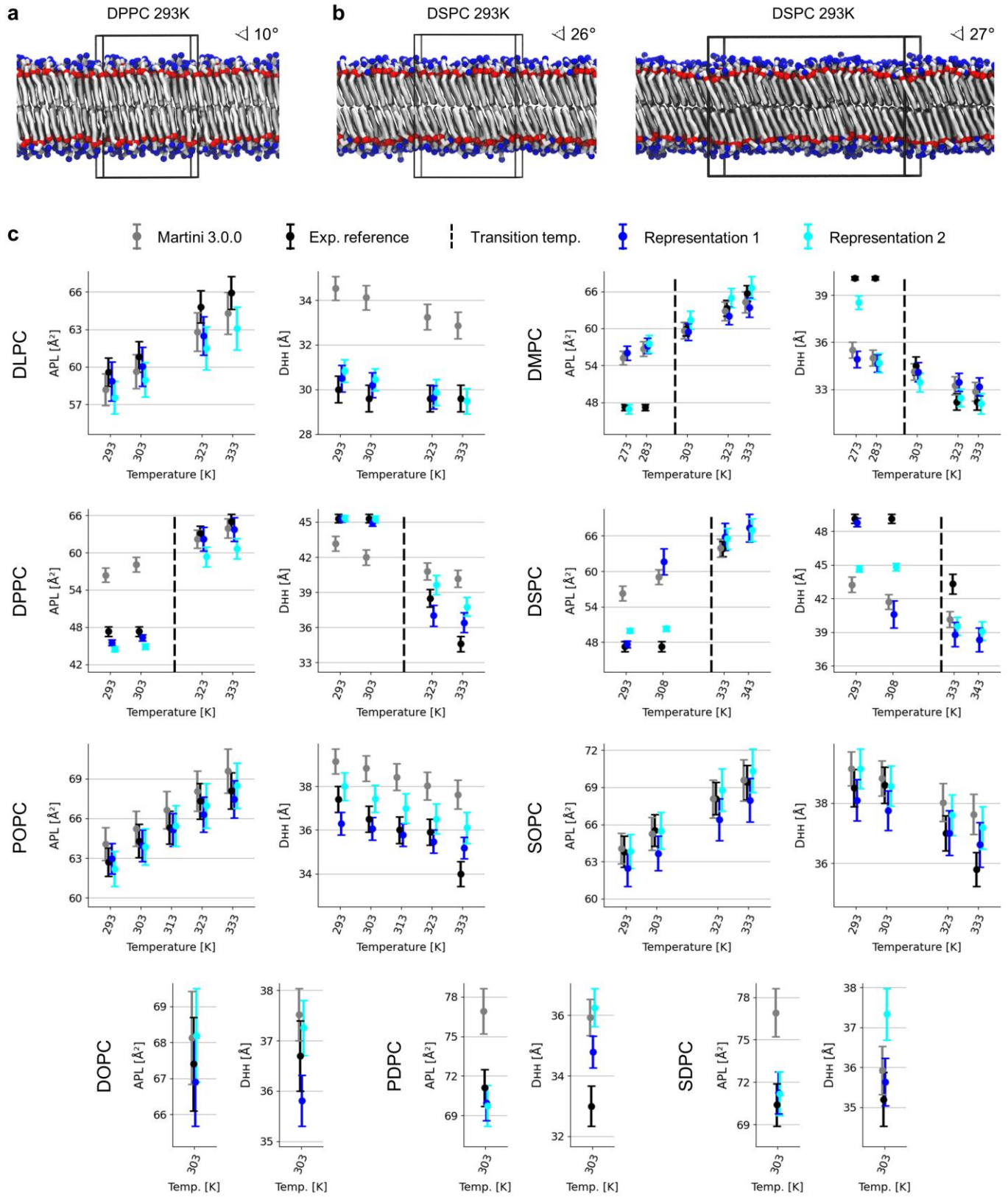

Figure 5. Summary of the APL and DHH thickness observed for the 8 PC lipid models in the training sets and additionally for SOPC, simulated in patches of lipid bilayers across multiple temperatures in the liquid and gel phases. (a) Snapshot of a DPPC bilayer simulated using Representation 2 and 128 lipids at 293K, exhibiting moderate tails tilting. (b) Snapshots of a DSPC bilayer simulated using Representation 2 at 293K, exhibiting significant tails tilting using either 128 lipids (left) or 512 lipids (right).

Lastly, we underline that experimental data collected for DMPC, DPPC and DSPC in the tilted gel phase ($L_{\beta'}$) correspond to measurements obtained from lamellar bilayers exhibiting non-negligible tilting of the tails with respect to the bilayer normal (~32 degrees)[20,30]. This tilting triggers a reduction of the hydrocarbon and D_{HH} thicknesses of ~13-15% in the $L_{\beta'}$ phase with respect to the straight gel phase (L_{β})[20,30]. Therefore, when evaluating the matching of bilayer thicknesses in between CG simulations and experimental measurements, the optimization protocol here implicitly formulates the requirement that models produce tilted tails for saturated lipids in the $L_{\beta'}$ phase. In practice, not all experimental objectives could be perfectly matched when optimizing the *bonded* parameters in the context of Representation 2 and we obtain a significant tilting of the tails only for DSPC in the $L_{\beta'}$ phase (Fig. 5b: ~26 degrees at 293K). Although the size of the molecular systems used during optimization is small, this tilt angle is stable also with increased system size (512 lipids) in 1 μs of production simulation (Fig. 5b). DMPC and DPPC bilayers exhibit only limited tails tilting (~10 degrees) when simulated up to 30K below their respective gel/liquid transition temperatures (Fig. 5a).

In Figure 5c we provide a summary of all APL and D_{HH} thickness measurements from extended simulations of lamellar bilayers (512 lipids each, 1 μs of production simulation) using the Martini 3.0.0[4] lipid models, together with the parameters obtained for Representations 1 and 2, along with reference experimental data[19,20,22,30]. We obtain improved parameters of both Representations 1 and 2 with respect to Martini PC lipids version 3.0.0 in the context of this training set, demonstrating the usefulness of such optimization protocols even when exploiting only the APL and D_{HH} thickness measurements during training. For Representation 1, however, the transition in between liquid and gel phases is not observed for DMPC and the transition temperature is not respected for DSPC (Fig. 5b). This can be associated to the usage of a mixture of small and regular beads in the lipid tails (Fig. S5). For Representation 2 the transition temperatures are better respected, while also the behavior of lipid mixtures appear improved, as a *posterior* result of the optimization of the structural properties in pure composition bilayer simulations.

IV. DISCUSSION

In this study, we apply the automated optimization strategy implemented in *SwarmCG*[25,26] for evaluating two putative refined CG representations of the lipid models in the framework of Martini, aiming at improving their thermodynamic properties. Because we optimize a finite number of *bonded* parameters in an informative context (*i.e.* rich training sets and *non-bonded* parameters remain constant, set to Martini 3.0.0[4]), we can eliminate uncertainties related to parameters tuning for focusing on evaluating the capabilities and limits of the CG representations (*i.e.* the choice of beads, their reference positioning, the topologies of the CG models and potentials used to describe *bonded* interactions). Our results indicate both Representations 1 and 2 allowed to improve upon the current version of the PC lipid models in Martini 3.0.0[4] (based on the 9 different lipid types included in the benchmark and on the objectives set) and the

protocol implemented in *SwarmCG*[25,26] is effective in this context for optimizing up to 77 *bonded* parameters (and likely more). Using mixed bead sizes in Representation 1 allowed to match the experimental objectives particularly in the liquid phases. The description of the phase behavior verified *a posteriori* for DOPC/DPPC mixtures appears relevant as well, but the DPPC model in this CG representation is a special case, that includes enough repeated big beads in the tails to enable the formation of a gel phase (which is not the case for DMPC and DSPC). Conclusively, the philosophy of Representation 1 cannot yield versatile building blocks for describing multiple types of lipids across phase states.

Instead, models obtained here with Representation 2 produce a more relevant modeling compromise, allowing to better approach all experimental objectives across phase states and producing improved phase transition temperatures with respect to the lipid models in Martini 3.0.0[4] (Fig. 5). In particular, the automated search of *bonded* parameters for the versatile building blocks of Representation 2 highlights that reproducing in Martini simulations the tilt angle of lipid bilayers in the $L_{\beta'}$ phase is possible. This is not trivial, as to our knowledge only one example of such CG simulations exists, where *bonded* parameters were manually optimized initially with the goal of obtaining a relevant modeling of the tilted ripple phase ($P_{\beta'}$) for DPPC[23] specifically, without considering the impact on the modeling of other lipid types (FF parameters are unavailable for this study). We envision that further working on the CG representation, potentially by slightly tweaking bead choices to further enhance the balance of *bonded* and *non-bonded* interactions in between lipid heads and tails, may allow to describe tails tilting in the $L_{\beta'}$ phase for relevant saturated lipids in the Martini framework.

As demonstrated, automated integrative modeling approaches such as *SwarmCG*[25,26] can be successfully leveraged not only for calibrating force field parameters, but also for (in)validating modeling philosophies, with higher throughput and enhanced certainty than what can be done manually. To this end, here we relied principally on experimental APL and D_{HH} thickness values for optimizing the *bonded* parameters of PC lipid models, iteratively simulating small patches of bilayers at a limited number of temperatures across phase states. We note, however, that this protocol could be extended with limited efforts to obtain *bonded* building blocks for a much larger set of lipid models in the context of Martini[4]. The training set can be further enriched by including additional bilayer systems (*e.g.* different lipid head types, sphingomyelins, *etc.*) and the loss function can directly evaluate also other experimental measurements providing transversal information (*e.g.* the hydrocarbon thickness would complement well the currently employed D_{HH} thickness and APL). Other available experimental properties such as bending modulus or diffusion constants may also be exploited at a slightly higher computational cost. Using annealing simulations directly in the training set could also enable finer estimation and tuning of the gel/liquid phase transition temperatures.

Further developments of *SwarmCG*[25,26] may also target other classes of molecules, such as DNA, peptides and proteins, which are well suited to the application of automatic parametrization approaches leveraging the cross-sampling of building blocks. Although the computational burden may seem important, such approaches scale well on HPC resources (*n.b.* in this study we only used 64 CPU cores). Their current limits

remain defined mostly by our ability to assemble rich training sets, including reliable experimental data to be used in concert with relevant CG representations, sufficiently descriptive of the degrees of freedom of the molecular systems of interest.

SUPPLEMENTARY MATERIAL

Supplementary material includes details on the functional form of the CG FF within which the parameters are optimized, the molecular models used in these experiments, their topologies, their optimized FF parameters obtained with *SwarmCG* in the context of Representations 1 and 2, as well as the implementation for usage with HPC resources. Additional details are also provided concerning the submolecular features observed in the CG models obtained at the end of the optimization experiments, underlining the relevance of the OT-B metrics.

ACKNOWLEDGEMENTS

G.M.P. acknowledges the funding received by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 818776 - DYNAPOL), by the Swiss National Science Foundation (SNSF grant: IZLIZ2_183336), and by H2020 under the FET Open RIA program (grant agreement no. 964386 - MIMICKEY). S.J.M. acknowledges ERC funding (grant agreement no. 101053661 - COMP-O-CELL). P.C.T.S. acknowledges the support of the French National Center for Scientific Research (CNRS) and the funding from a research collaboration agreement with PharmCADD. K.B.P. acknowledges funding through FTP (Grant: 0136-00148B) and NNF-ROBUST (Grant: NNF18OC0032608). The authors also acknowledge the computational resources provided by the Swiss National Supercomputing Center (CSCS).

AUTHOR DECLARATIONS

Conflicts of Interest

The authors have no conflicts of interest to disclose.

Author Contributions

C.E.M. devised and implemented the algorithm. C.E.M and K.B.P. wrote the paper. C.E.M., K.B.P., R.C., M.P. and C.C. assembled the benchmark and performed the experiments. M.C. performed the LENS analysis. All authors discussed the molecular models and modeling philosophies. G.M.P., P.C.T.S. and S.J.M. supervised the work. All authors agreed on the final form of the paper.

DATA AVAILABILITY

The code used in this study for the optimization of the CG lipid models is available at: https://github.com/GMPavanLab/SwarmCGM, together with the models obtained at convergence for each experiment, as well as the configuration files used in this study and all material necessary for reproducing the results.

ABBREVIATIONS

DLPC: 1,2-dilauroyl-sn-glycero-3-phosphocholine

DMPC: 1,2-dimyristoyl-sn-glycero-3-phosphocholine

DPPC: 1,2-dipalmitoyl-sn-glycero-3-phosphocholine

DSPC: 1,2-distearoyl-sn-glycero-3-phosphocholine

POPC: 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine Δ9-Cis

SOPC: 1-stearoyl-2-oleoyl-sn-glycero-3-phosphocholine Δ9-Cis

DOPC: 1,2-dioleoyl-sn-glycero-3-phosphocholine Δ9-Cis

PDPC: 1-palmitoyl-2-docosahexaenoyl-sn-glycero-3-phosphocholine Δ4,7,10,13,16,19-Cis

SDPC: 1-stearoyl-2-docosahexaenoyl-sn-glycero-3-phosphocholine Δ4,7,10,13,16,19-Cis

DLPE: 1,2-dilauroyl-sn-glycero-3-phosphoethanolamine

DMPE: 1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine

DLPS: 1,2-dilauroyl-sn-glycero-3-phospho-L-serine

DMPS: 1,2-dimyristoyl-sn-glycero-3-phospho-L-serine

REFERENCES

1. Jin, J., Pak, A. J., Durumeric, A. E. P., Loose, T. D. & Voth, G. A. Bottom-up Coarse-Graining: Principles and Perspectives. *J. Chem. Theory Comput.* **18**, 5759–5791 (2022).
2. Marrink, S. J. *et al.* Two decades of Martini: Better beads, broader scope. *WIREs Comput. Mol. Sci.* **13**, e1620 (2023).
3. Marrink, S. J., de Vries, A. H. & Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **108**, 750–760 (2004).
4. Souza, P. C. T. *et al.* Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nat. Methods* **18**, 382–388 (2021).
5. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).

6. Abellón-Ruiz, J. *et al.* Structural basis for maintenance of bacterial outer membrane lipid asymmetry. *Nat. Microbiol.* **2**, 1616–1623 (2017).
7. Van Eerden, F. J., Melo, M. N., Frederix, P. W. J. M., Periole, X. & Marrink, S. J. Exchange pathways of plastoquinone and plastoquinol in the photosystem II complex. *Nat. Commun.* **8**, 15214 (2017).
8. Yen, H.-Y. *et al.* PtdIns(4,5)P 2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **559**, 423–427 (2018).
9. Heidenreich, M. *et al.* Designer protein assemblies with tunable phase diagrams in living cells. *Nat. Chem. Biol.* **16**, 939–945 (2020).
10. Vögele, M., Köfinger, J. & Hummer, G. Hydrodynamics of Diffusion in Lipid Membrane Simulations. *Phys. Rev. Lett.* **120**, 268104 (2018).
11. D'Agostino, M., Risselada, H. J., Lürick, A., Ungermann, C. & Mayer, A. A tethering complex drives the terminal stage of SNARE-dependent membrane fusion. *Nature* **551**, 634–638 (2017).
12. Souza, P. C. T. *et al.* Protein–ligand binding with the coarse-grained Martini model. *Nat. Commun.* **11**, 3714 (2020).
13. Alessandri, R. *et al.* Martini 3 Coarse-Grained Force Field: Small Molecules. *Adv. Theory Simul.* **5**, 2100391 (2022).
14. Vainikka, P., Thallmair, S., Souza, P. C. T. & Marrink, S. J. Martini 3 Coarse-Grained Model for Type III Deep Eutectic Solvents: Thermodynamic, Structural, and Extraction Properties. *ACS Sustain. Chem. Eng.* **9**, 17338–17350 (2021).
15. Pezeshkian, W. *et al.* Molecular architecture and dynamics of SARS-CoV-2 envelope by integrative modeling. *Structure* (2023) doi:10.1016/j.str.2023.02.006.
16. Stevens, J. A. *et al.* Molecular dynamics simulation of an entire cell. *Front. Chem.* **11**, (2023).
17. Hoffmann, C., Centi, A., Menichetti, R. & Bereau, T. Molecular dynamics trajectories for 630 coarse-grained drug-membrane permeations. *Sci. Data* **7**, 51 (2020).
18. Marrink, S. J. & Mark, A. E. Molecular Dynamics Simulation of the Formation, Structure, and Dynamics of Small Phospholipid Vesicles. *J. Am. Chem. Soc.* **125**, 15233–15242 (2003).
19. Kučerka, N., Nieh, M.-P. & Katsaras, J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. Acta BBA - Biomembr.* **1808**, 2761–2771 (2011).

20. Tristram-Nagle, S., Liu, Y., Legleiter, J. & Nagle, J. F. Structure of Gel Phase DMPC Determined by X-Ray Diffraction. *Biophys. J.* **83**, 3324–3335 (2002).
21. Jarin, Z., Newhouse, J. & Voth, G. A. Coarse-Grained Force Fields from the Perspective of Statistical Mechanics: Better Understanding of the Origins of a MARTINI Hangover. *J. Chem. Theory Comput.* **17**, 1170–1180 (2021).
22. Marquardt, D. *et al.* The structures of polyunsaturated lipid bilayers by joint refinement of neutron and X-ray scattering data. *Chem. Phys. Lipids* **229**, 104892 (2020).
23. Sharma, P., Desikan, R. & Ayappa, K. G. Evaluating Coarse-Grained MARTINI Force-Fields for Capturing the Ripple Phase of Lipid Membranes. *J. Phys. Chem. B* **125**, 6587–6599 (2021).
24. Carpenter, T. S. *et al.* Capturing Phase Behavior of Ternary Lipid Mixtures with a Refined Martini Coarse-Grained Force Field. *J. Chem. Theory Comput.* **14**, 6050–6062 (2018).
25. Empereur-Mot, C. *et al.* Swarm-CG: Automatic Parametrization of Bonded Terms in MARTINI-Based Coarse-Grained Models of Simple to Complex Molecules via Fuzzy Self-Tuning Particle Swarm Optimization. *ACS Omega* **5**, 32823–32843 (2020).
26. Empereur-mot, C. *et al.* Automatic multi-objective optimization of coarse-grained lipid force fields using SwarmCG. *J. Chem. Phys.* **156**, 024801 (2022).
27. Uppamoochikkal, P., Tristram-Nagle, S. & Nagle, J. F. Orientation of Tie-Lines in the Phase Diagram of DOPC/DPPC/Cholesterol Model Biomembranes. *Langmuir* **26**, 17363–17368 (2010).
28. Nobile, M. S. *et al.* Fuzzy Self-Tuning PSO: A settings-free algorithm for global optimization. *Swarm Evol. Comput.* **39**, 70–85 (2018).
29. Sengupta, S., Basak, S. & Peters, R. A. Particle Swarm Optimization: A Survey of Historical and Recent Developments with Hybridization Perspectives. *Mach. Learn. Knowl. Extr.* **1**, 157–191 (2019).
30. Nagle, J. F., Cognet, P., Dupuy, F. G. & Tristram-Nagle, S. Structure of gel phase DPPC determined by X-ray diffraction. *Chem. Phys. Lipids* **218**, 168–177 (2019).
31. Pele, O. & Werman, M. Fast and robust Earth Mover's Distances. in *2009 IEEE 12th International Conference on Computer Vision* 460–467 (2009).
32. Rubner, Y., Tomasi, C. & Guibas, L. J. The Earth Mover's Distance as a Metric for Image Retrieval. *Int. J. Comput. Vis.* **40**, 99–121 (2000).
33. Villani, C. *Optimal Transport: Old and New*. (Springer-Verlag, 2009).

34. Pele, O. & Werman, M. A Linear Time Histogram Metric for Improved SIFT Matching. in *Computer Vision – ECCV 2008* (eds. Forsyth, D., Torr, P. & Zisserman, A.) 495–508 (Springer, 2008).
35. Crippa, M., Cardellini, A., Caruso, C. & Pavan, G. M. Detecting dynamic domains and local fluctuations in complex molecular systems via timelapse neighbors shuffling. Preprint at https://doi.org/10.48550/arXiv.2212.12694 (2022).