A Data-Driven Dimensionality Reduction Approach to Compare and Classify Lipid Force Fields

Riccardo Capelli,^{*,†} Andrea Gardin,[†] Charly Empereur-mot,[‡] Giovanni Doni,[‡]
and Giovanni M. Pavan^{*,†,‡}

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

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

E-mail: riccardo.capelli@polito.it; giovanni.pavan@polito.it

Abstract

Molecular dynamics simulations of all-atom and coarse-grained lipid bilayer models are increasingly used to obtain insights useful for understanding the structural dynamics of these assemblies. In this context, one crucial point concerns the comparison of the performance and accuracy of classical force fields (FFs), which sometimes remains elusive. To date, the assessments performed on different classical potentials are mostly based on the comparison with experimental observables, which typically regard average properties. However, local differences of structure and dynamics, which are poorly captured by average measurements, can make a difference, but these are non-trivial to catch. Here we propose an agnostic way to compare different FFs at different resolutions (atomistic, united-atom, and coarse-grained), by means of a high-dimensional similarity metrics built on the framework of Smooth Overlap of Atomic Positions (SOAP). We compare and classify a set of 13 force fields, modeling 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayers. Our SOAP kernels-based metrics allows us to compare, discriminate and correlate different force fields at different model resolutions in an unbiased, high-dimensional way. This also captures differences between FFs in modeling non-average events (originating from local transitions), such as for example the liquid-to-gel phase transition in dipalmitoylphosphatidylcholine (DPPC) bilayers, for which our metrics allows to identify nucleation centers for the phase transition, highlighting some intrinsic resolution limitations in implicit *vs.* explicit solvent force fields.

Graphical TOC Entry

Keywords

Lipid bilayers, Coarse Graining, SOAP, Dimensionality Reduction, Unsupervised Clustering

Introduction

Lipid membranes are ubiquitous in biological systems, and their chemical and mechanical characteristics directly impact the regulation of the cell machinery.¹ Membranes constitute a barrier between the cell and its external environment, as well as define different structures and organelles within the cell; they are involved in transport processes,² signaling,³ and protein interactions,⁴ to name a few. A plethora of experimental techniques like NMR,⁵ Calorimetry,⁶ SANS⁷ and SAXS⁸ have been applied to lipid bilayers to obtain average structural and dynamic information at different resolutions. This large amount of experimental data paved the way to the creation and the cross-validation of reliable models that can be simulated by means of molecular dynamics (MD). The use of all atom (AA) provides, in principle, an atomic-resolution computational microscope for investigating the dynamic evolution of membranes in great detail.⁹ Despite the tremendous advance in computational capabilities observed in the last decades, classical MD at atomistic resolution is still unable to cover all the time scales of biological interest.¹⁰ For this reason, starting from the beginning of the 90s, various models with a reduced number of degrees of freedom were proposed, from the united atom (UA) representations, where the aliphatic hydrogen atoms are removed and their mass is added to the bound heavy atom,¹¹ to coarse-grained (CG), where a single “CG bead” formed by usually 2-5 heavy atoms,¹² to super-CG models, where a single lipid can be represented by 3-4 larger CG beads.¹³ The reduction of the number of degrees of freedom provides a dramatic speed up in the simulations, which is nonetheless accompanied by an unavoidable loss of entropic contribution (and thus accuracy), typically compensated by properly compensating the enthalpic contributions.¹⁰ The evaluation of the precision and the performance of a force field (at any level of resolution) is in general performed comparing average equilibrium observables obtained from simulations of the bilayer models to the experimentally available ones. A large body of work has addressed the problem of comparing state-of-the-art lipid force fields (FFs) and their accuracy with respect to increasingly precise experimental data.^{14–17} However, the general assessment of the performance of a FF remains

a difficult problem which concerns multiple parameters at the same time. Moreover, comparing various FFs becomes particularly awkward when different representations/resolutions in the modeling of the chemical system are employed, and when the internal dynamic organization of the membrane, its uniformity/non-uniformity and (local) dynamic fluctuations become important.

In this work, we consider 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC, Fig. 1) bilayers as a reference system to compare a set of 13 different FFs at various levels of resolution. In particular, the bilayer models have been simulated for 1 μ s of MD at 303 K in the NPT ensemble using different molecular representations. As summarized in Table 1, we compared 3 AA models (Slipids,^{18,19} Charmm36,²⁰ and AMBER LIPID17²¹), 3 UA models (Berger Lipid FF,²² GROMOS43a1-s3,²³ and GROMOS-CKP²⁴), as well as 3 wet CG (Martini 2.2,^{25,26} Martini 3.0 beta 3.2, and Sirah FF²⁷), 2 polarizable CG (Martini 2.2p^{28,29} and Martini 2.3p³⁰), and 2 implicit-solvent CG models (Dry Martini,³¹ with the original POPC mapping and the new mapping). Our analysis demonstrates that state of the art FF potentials exhibit differences in terms of widely studied equilibrium and dynamic average observables (Fig. 2), making both the comparison (between FFs at the same level of detail), and the assessment of the accuracy of the coarse-graining protocol (thus, between FFs with different CG levels) not trivial to perform in a rigorous, unbiased and unambiguous way. Recently, a large number of dimensionality reduction/machine-learning techniques emerged as powerful tools for evaluating the collective dynamics behavior of complex chemical/molecular systems in equilibrium and non-equilibrium conditions.^{32–40} Among those, Smooth Overlap of Atomic Position (SOAP) vectors³⁵ are extremely useful to provide a high-dimensional, agnostic and rich description of molecular environments in molecular systems. SOAP has been successfully applied in exploring the conformational landscape of single molecules,³⁶ nucleation phenomena,^{34,37} molecular assemblies classification,³⁸ as well as the formation, stability and intrinsic dynamic complexity of soft supramolecular assemblies,^{39,40} providing

a rich structural/dynamical characterization of complex molecular systems that is not easy to obtain with human-based approaches.

Herein, we employ a distance between the average SOAP representations to compare and classify the different data ensembles obtained from the MD simulations using the different FFs. The analyzed FFs have some internal differences in their organization that can be detected with SOAP: this allow us to classify the FFs based on a global score of similarity. Our results with the SOAP metrics are validated by comparing them with an analogous analysis based on a well-known distribution metrics (i.e., the Jensen-Shannon divergence⁴¹), finding a substantial agreement between the two.

Finally, to provide further evidence of the local chemical-physical origins of differences between the FFs that the SOAP environment representation allows to elucidate, we focused on the comparison between two widely employed explicit-solvent and implicit-solvent CG models (namely, Martini 2.2,^{25,26} and Dry Martini³¹). We investigated the liquid-to-gel phase transition in a pure bilayer of dipalmitoylphosphatidylcholine (DPPC) lipids simulated at different temperatures: in gel phase, mixed gel-liquid, and liquid phase. By computing individual SOAP vectors for each lipid molecule in the bilayer across MD simulations at different temperatures, and clustering these with Probabilistic Analysis of Molecular Motifs (PAMM),⁴² we could clearly identify nucleation centers underpinning the phase transition, comparing how well CG force fields at different scales (wet and dry) reproduce the features of the system.

Results and Discussion

Comparing Lipid Force Fields Using Average Observables

We started by simulating a bilayer formed by 128 POPC molecules (64 per leaflet) at 303 K using 3 different FFs resolutions (Fig. 1): AA (CHARMM36, Slipids, LIPID17), UA

(Berger, Gromos-CKP, Gromos-43a1-s3) and CG (Martini 2.2, Martini 3.0 beta 3.2, Sirah 2.1 and Dry Martini in its 2014 and 2016 versions for POPC). A summary of the selected models is provided in Table 1. For all the FFs, except the UA and Sirah 2.1, we generated the initial conformation for atomistic and Martini bilayers using the online tool CHARMM-GUI,^{43,44} which have been then minimized and equilibrated according to well established protocols (for details see Supporting Information, SI). For the UA FFs, we obtained the input files for Berger lipids from the internet,^{45,46} and for Gromos-CKP and Gromos43a1-s3 from Lipidbook,⁴⁷ following also in this case the minimization and equilibration protocols indicated by CHARMM-GUI. We performed 1 μ s of MD simulation using GROMACS 2018.6⁴⁸ patched with PLUMED^{49,50} (details on the simulation parameters used for each system are available in the SI). For Sirah we employed AMBERTools using the simulation parameters given on the home page of the FF (<http://www.sirahff.com>, see also SI). Also in this case, we performed a 1 μ s-long MD simulation for data production.

Table 1: Summary of the lipid Force Fields compared herein. For Dry Martini we employed 2 different available mappings for POPC: the one presented in the original version of the FF, that we named Dry Martini (2014), and a second mapping made available in 2016, which added one CG bead in the unsaturated chain, that we named Dry Martini (2016).

Force Field	Ref.	Resolution	Year
Slipids	18,19	All-atom	2012
CHARMM36	20	All-atom	2010
AMBER LIPID17	21	All-atom	2018
Berger	22	United Atom	1997
GROMOS43a1-s3	23	United Atom	2009
GROMOS-CKP	24	United Atom	2012
Martini 2.2	25,26	Wet CG	2007
Martini 3.0 beta 3.2	—	Wet CG	2018
Sirah 2.1	27	Wet CG	2019
Martini 2.2p	28,29	Polarizable CG	2013
Martini 2.3p	28–30	Polarizable CG	2020
Dry Martini (2014)	31	Dry CG	2014
Dry Martini (2016)	31	Dry CG	2016

Initially focusing on the equilibrium structural properties of the lipid models, we extracted from the production phase MD simulations the area per lipid (A_L), bilayer thickness (D_{HH})

Figure 1: Representation of POPC. (a) Chemical structure of a POPC molecule, (b) the 5 different mappings employed in our study.

and phosphate group density profiles for each case (see Fig.2). From the experimental point of view, all these properties depend to some extent on the technique used in the measurement, displaying large deviations in the reported values.⁵¹ In the A_L evaluation (top left in Fig. 2), we can see a general consensus between all FFs, which fall in the region of one of the two experimental A_L estimations available in the literature^{52,53} (estimations performed using a combination of SANS/SAXS measurements). Regarding the D_{HH} (top center in Fig. 2), in general all compared CG models overestimate the experimentally available values (with a maximum deviation of $\sim 20\%$ for Dry Martini 2014). The density profiles of the phosphate groups show a large variability both in height and in width (top right in Fig. 2). This can be likely imputed to differences in the CG beads, and in the bead-bead interactions, used in the different CG FFs - slightly different approximations in the enthalpy/entropy balance in the models (that unavoidably accompany the various CG schemes⁵⁴) can for example make the thermal vibrations larger/smaller, broadening/narrowing the density distributions. A great variability is present in membrane compressibility (bottom left in Fig. 2), irrespectively of the FF resolution. For diffusion coefficient (bottom center in Fig.2), we see a general consensus between the AA and UA FFs and the Sirah 2.1 CG FF. A higher diffusion coefficients

are obtained for the CG FFs belonging to the Martini family, which can be nonetheless expected for CG FFs due to the enhanced sampling guaranteed by the CG scheme. It is worth underlining that a direct comparison of diffusion coefficients between different resolution FFs may be not meaningful, as the sampling may be different. Nonetheless, it is interesting to note that a CG FF like Sirah shows a diffusion coefficient that is closed to those of AA/UA FFs rather than to those of the other CG FFs. This is a clear evidence that different approaches in the CG parametrization may result in a different sampling of the models.

From this initial analysis, the performances of every FF are dramatically dependent on the observables considered.

Figure 2: Comparison of membrane observables obtained in this study for all the analyzed FFs. We computed area per lipid (A_L , top left), membrane thickness (D_{HH} , top center), density profiles for the phosphate groups (top right), membrane compressibility (bottom left), and diffusion coefficient (bottom center). Diffusion coefficient calculations are not corrected for PBC effects (see SI).

A SOAP Metrics to Compare Lipid Force Fields

While such parameters may be useful to compare between the different FFs, it is not easy to obtain a complete, exhaustive picture from such a low-dimensional analysis. A more general, high-dimensional analysis would be desirable for a more rigorous comparison. To define a metrics that can be applied to compare lipid models having a different resolution (AA, UA and CG), we started from considering a unified lipid representation well suited for such comparison. We chose 4 sites (beads) in the lipids structure that are in common between all considered lipid models, and that we use for our SOAP analysis: one bead for the lipid charged head group, one bead for the phosphate group, and one bead for each of the two alkylic lipid tails. In particular, one SOAP vector is centered in the phosphate bead of each lipid molecule in the system, and it considers in the analysis all other beads in the atomic (particle) environment that surrounds each lipid (phosphate) center in the membrane. Thus, we obtain a SOAP power spectrum that is indicative of the local environment that surrounds each phosphate group in the bilayer, which accounts of e.g. levels of order/disorder, spatial displacements and packing of the lipids heads, phosphate groups and the lipid tails. In this way, we come out with a SOAP power spectrum for every lipid in the lipid bilayer model, at every MD snapshot. Averaging the power spectra calculated for all lipids in the model at a given MD snapshot, we obtain an indication of the average environment that surrounds the lipids in the membrane model (average “static” lipids SOAP power spectrum). Finally, averaging these on all the equilibrium MD snapshots considered in the analysis (we considered 100 MD snapshots, one frame every 10 ns, extracted from the equilibrated phase MD trajectories), we finally obtain a global insight into the complexity of the lipid bilayer in terms of the structural and dynamic features of the local environment surrounding, on average, the lipid centers in the system in the equilibrium MD regime. Breaking down the SOAP analysis in this way has three main advantages. (i) The SOAP calculations for the systems are computationally manageable. (ii) This helps in focusing the analysis on the supramolecular dynamics of the lipid bilayers (that is, the reshuffling, movements,

order/disorder in the displacement of the lipids in the membrane), reducing noise that is intrinsically present in the MD trajectories. Finally, (iii) this guarantees to compare the different models studied herein using a common metrics. Assuming that we are sampling well the conformational space in the equilibrium MD regime for each simulated system, we therefore obtained 13 average SOAP spectra that are informative on the structural dynamics of POPC lipid bilayers at 303 K.

From these average spectra, we could then compute the SOAP distances between each of the FFs (details in Computational Methods and SI), which provides a distance matrix which is indicative of the similarity (in the SOAP space) between the FFs (Figure 3a).

Figure 3: Force fields Distance matrix and 2D representation. a) Reciprocal SOAP distance matrix for all the 13 FFs under comparison. A hierarchical clustering allows to group them and to highlight similarities/difference between them. Lighter blue colors indicate shorter SOAP distance between FFs, meaning higher similarity between the FFs. b) Compression to a 2D representation of the high-dimensional distances obtained from the distance matrix of Fig. 3a (calculated on an N-dimensional space) using a multidimensional scaling (MDS) algorithm.⁵⁵ The cluster formed by wet CG Martini FFs and the significant improvement of the new mapping in the Dry Martini FF for POPC appears evident. We also highlighted the two example distances between Sirah 2.1 and the two versions of Dry Martini CG FFs. The color of the points that identify the various FFs is chosen to divide the various FF families based on their resolution: red for AA, blue for UA, green for “wet” Martini, brown for Sirah, and gray for Dry Martini FFs.

Figure 3a shows the obtained distance matrix and the hierarchical clustering for the performance of the individual FFs put in comparison one another. In particular, light cyan colors identify short average SOAP distances between the FFs, indicating that these FFs behave similarly. Dark colors identify larger SOAP distances, and increased discrepancy between the FFs behavior. From Figure 3a, it appears evident that all wet and polarizable CG Martini FFs represent the lipid behavior in a similar way (green symbols in the matrix are united by light cyan colors). In particular, it is worth noting that Martini 2.2p and Martini 2.3p CG FFs are superimposed, because the representation of POPC is identical in these two FFs. Another interesting point is the improvement of the Dry Martini performances between the initial 2014 version/mapping (light grey) and the 2016 one (dark grey). The hierarchical adjacency graph of Figure 3a shows that this modification makes the newer version of this CG FF closer to the wet CG Martini FFs, and to the AA and UA FFs, respect to the older version. Regarding AA and UA FFs, we observe a substantial proximity between all of them in the SOAP metrics space. Interestingly, the Berger FF (cyan), which is the oldest developed model in this comparison (1997), shows a reduced SOAP distance from 2 recent AA FFs: Slipids (2012) and AMBER LIPID17 (2017).

Such a use of distance measurements between the average SOAP environments has been validated by calculating the Jensen-Shannon divergence⁴¹ between the probability distribution sampled via MD simulations for all FF pairs (details in the methods section in the SI). This test has been carried out to verify the degree of agreement between the two approaches. A correlation plot between the SOAP distance and the Jensen-Shannon divergence computed between the average SOAP spectra is available in the SI (Figure S1). The distance metrics calculated between the average SOAP spectra then allowed us to discriminate and compare between different FFs for POPC at a temperature of 303 K, for which all systems are in the liquid phase.

Capturing Gel-Liquid Phase Transitions on a Local Level with SOAP

While this approach allows us to compare the FFs between them in a rather comprehensive way (structure, dynamics, order, etc.) - we learn for example that some CG FFs are closer to UA and AA FFs compared to other ones -, it is not straightforward to link such information extracted from high-dimensional analyses to human-comprehensible data. At the same time, the comparison with a few experimental data (Figure 2a) does not suffice to obtain a clear and comprehensive picture. For example: what is the difference between Martini 2.2 and Dry Martini 2016 CG FFs in the modeling of a lipid bilayer? Comparing the data of Figure 2a would suggest that these two CG FF classes behave rather similarly (apart from the compressibility). These have also a similar SOAP distances from e.g. AA FFs (Figure 3a, red colors). However, in Figure 3b these also appear as equally distant between them. Recently, it was shown that Martini 2.2 and Dry Martini CG FFs respectively tend on average to over-structure and under-structure the bilayers compared to AA FFs.⁵⁴ However, all these analyses and comparisons provide evidences that are limited to the average characterization of the bilayers, while, on the other hands, it has been shown that the behavior of complex supramolecular assemblies (such as also lipid bilayers) may be strongly controlled by local events, or fluctuations, that cannot be captured with average evaluations.^{39,40,56}

To move our investigation to a deeper level, similarly to a recent study³⁷ where typical ice nucleation sites were probed in a QM-based liquid water model by means of a SOAP analysis, we investigated the transition of a lipid bilayer between the liquid phase and the gel phase. In particular, we were interested in using a high-dimensional SOAP-based approach to detect and characterize local nucleation processes underpinning the formation a new phase in the lipid bilayers during a phase transition (or, in the absence of nucleation events, to identify limitations in the FF representation). The experimental melting temperature T_m for POPC is approximately 273 K,⁵⁷ which is inconvenient for working with classical MD. We thus decided to turn to another well-studied lipid, dipalmitoylphosphatidylcholine (DPPC), which transition from the gel to liquid phase occurs at a temperature of ~ 315 K.⁵⁸ For the

wet Martini FF, T_{melting} has been reported in semi-quantitative agreement at $\sim 295 \pm 5$ K,⁵⁹ while for Dry Martini T_{melting} has been estimated at ~ 333 K.⁶⁰ We built a bilayer model composed of 1152 DPPC molecules (576 per leaflet), parametrized using the Martini 2.2 and Dry Martini CG FFs (for DPPC there is no difference, as for POPC, between 2014 and 2016 versions of Dry Martini). Both CG models were minimized and equilibrated using the same MD procedure as in the POPC case (see SI for details). In particular, the lipid bilayers were simulated at 273 K, 293 K, 323 K, 333 K and 353 K for both wet Martini 2.2 and Dry Martini CG models. All the production MD simulations were 1 μ s-long.

We used again 4 beads per-lipid for the SOAP analysis (one bead for choline group, one for the phosphate group, and one bead for the last alkylic group of each tail in the lipids), and we set the phosphate bead as the SOAP center for calculating each spectrum. This representation is a compromise between computational cost and comprehensiveness of the analysis, and it is particularly well suited for studying phase transitions, as the tails arrangement is indicative of the phase transition in a lipid bilayer. In this analysis, we computed the SOAP spectra independently for each single lipid in the bilayer without averaging them. We thus obtained 1152 spectra per-analyzed frame (one every 10 ns of MD). We then used an unsupervised clustering method, the Probabilistic Analysis of Molecular Motifs (PAMM),⁴² to classify the lipids based on their surrounding environment, their local levels of order and disorder, and on the fluctuation/persistence of the latter along the equilibrium MD trajectories. This allowed us to classify the lipids in the bilayer during the simulation based on their SOAP spectra, and to discriminate those belonging to the gel or to the liquid phase. The clustering procedure was performed on both the FFs merging all the spectra obtained from the 5 different temperatures (details in Computational Methods), to identify the same clusters along all the temperatures. In particular, from those lipids that dynamically change cluster during the equilibrium MD, we could obtain a rather comprehensive dynamic picture of the equilibrium between gel and liquid phases at the different temperatures, and to compare how these are reproduced in an explicit-solvent or in an implicit-solvent CG FF. The

results of the PAMM analysis at different temperatures for the Martini 2.2 and Dry Martini DPPC bilayers are reported in Figure 4.

Figure 4: Unsupervised clustering (PAMM) analysis of lipid domains in the DPPC bilayer simulations. Every column of the figure refers to a simulation temperature (273 K, 293 K, 323 K, 333 K, 353 K). The top half of the figure shows the results for Martini 2.2, while the bottom half reports the results obtained using the Dry Martini model. For every half, we report (top) the projection of the SOAP spectra of the lipids obtained along the MD trajectory on the 2 first PCA components colored based on the assigned cluster (red: gel phase, blue: liquid phase), example MD snapshot of the membrane from the lateral and top view with the cluster coloring, showing the localization/distribution of the lipids belonging to the different phases (middle), and the transition matrices reporting the transition probabilities of the lipids between the clusters (bottom). A phase transition between 293-323 K is evident in the Martini 2.2 model. Despite a growth in the ordered (gel) cluster population (red cluster), the same abrupt transition is not observed using the Dry Martini model.

In the case of Martini 2.2 (top half in Figure 4), the PAMM analysis is able to discriminate between lipids belonging to an ordered phase (red cluster) and a disordered phase (blue cluster), which can respectively interpreted as the gel and the liquid phases in the lipid bilayer. As expected, the relative populations within the 2 clusters shows an inversion from $T < T_{\text{melting}}$ to $T > T_{\text{melting}}$ (Figure 5). Below 293 K the system is dominated by red lipids, while above 323 K by blue ones. We can clearly observe nucleation events along the MD trajectory at $T \simeq T_{\text{melting}}$ (see snapshot at 293 K in Figure 4). While dynamical data extracted from approximated CG models must be considered qualitatively, the data in the transition matrices (obtained using a timestep between the analyzed MD snapshots of 10 ns) provide a comprehensive picture of the intrinsic dynamics of the systems. For example, at 293 K from the first row of the matrix (normalized to sum to 1), we can evince that in equilibrium condition in the bilayer a red lipid (gel) remains red ~ 92 % of the sampled MD trajectory, and with an ~ 8 % probability undergoes transition into blue (liquid). Interestingly, the blue (liquid) cluster is stable only at ~ 44 %, while the probability for blue lipids to undergo transition into red is even higher (~ 56 %). This shows that we are still below the transition temperature to liquid. In fact, at 323 K the situation is completely reversed, and from the statistical point of view the liquid state for the lipids become dominant. This analysis highlights the (dynamic) equilibrium present between the phases (clusters) in the system, and allows to observe on a local perspective how this is perturbed/moved while changing the temperature.

While the phase transition, and its local origin, appear to be well-reproduced by the explicit-solvent Martini 2.2. FF, the implicit-solvent Dry Martini FF provides a different picture (lower half in Figure 4). Also in this case, the PAMM analysis identified the same two clusters (red lipids in gel phase, and blue lipids in liquid phase). However, although the gel (red) population remains somehow inversely proportional to T , we cannot observe a sharp phase transition in the simulations performed across the temperatures. Shown in Figure 5, the plot of the clusters populations at the different temperatures highlights a transition

Figure 5: Cluster populations for Martini 2.2 (top) and Dry Martini (bottom) DPPC bilayers in function of simulation temperature. In the case of Martini 2.2 a gel (red) to liquid (blue) phase transition appears evident. Conversely, for Dry Martini this kind of transition is absent.

point between 293-323 K for Martini 2.2, while this is absent for Dry Martini. In particular, in the Dry Martini FF the amount of lipids that belong to gel phase increases in number while decreasing the temperature, but these do not become dominant. Basically, the bilayer remains always liquid, although this becomes on average more static lowering the temperature. This is probably due to the unavoidable approximations resulting from encoding both solute-solute and solute-solvent interactions (in the explicit-solvent model) into solute-solute equivalent interactions in the implicit-solvent model. This highlights the role of having explicit solvent molecules in the system (even a very minimalistic representation thereof, as in the Martini scheme) in reproducing locally-triggered events that are poorly reproduced when the effect of the solvent is averaged in the system. Also, our results indicate that the main difference between Martini 2.2 and Dry Martini (indicated by the average SOAP distance of Figure 2b) is local - namely, in how these two FFs model the local lipid environments in the bilayers. This demonstrates how bottom-up (e.g. comparison with the AA models) or top-down (comparison with average experimental data) approaches comparing FFs simply relying on average data on the entire/global bilayers may be insufficient. This also shows the potential of high-dimensional data-driven analyses in providing detailed information in this sense.

Conclusions

In conclusion, we presented a data-driven dimensionality reduction approach providing a metrics based on SOAP spectra that is able to quantify the similarity between force fields at different levels of resolution. We applied this to compare how various force fields (FFs) model lipid bilayers at different levels of resolution (all atom, AA, united atom, UA, or coarse-grained, CG). Using POPC lipids as a reference case, our analysis highlights the good agreement between state of the art AA and UA FFs. Regarding wet CG FFs, we observed

a substantial equivalence in the wet Martini family, and a significant improvement in the global representation of POPC bilayers in the 2016 version of Dry Martini compared to the original Dry Martini version (2014), which is found closer to the other wet CG FF analyzed, Sirah 2.1, and to the fine-grained FF studied. These analyses offer great opportunities to obtain detailed insight into the different representation of the model bilayers by the different FFs which are poorly captured by conventional, average analyses. Our local SOAP-PAMM analysis, for example, allows to identify in an unbiased and unsupervised way the different states for the lipids in a DPPC bilayer at different temperatures, below and above the transition temperature T_{melting} . In this way, we can clearly identify the lipids belonging to the ordered (gel) or to the disordered (liquid) phases in the bilayers, and we can reconstruct the equilibrium structural dynamics inside the bilayers, as well as the dynamic nucleation of gel/liquid phases across the temperatures in a comprehensive way. Comparing an explicit-solvent and an implicit-solvent CG FFs (Martini 2.2 and Dry Martini), our results clearly demonstrate that while Martini 2.2 model well the phase transition, the implicit-solvent Dry Martini FF does not model well transitions that are strongly triggered by local events. A next possible step to further deepen the investigation exploiting this SOAP metrics could be the specific comparison between AA/UA force fields, e.g. considering a larger number of beads in the lipids molecules as particles accounted in the SOAP analysis. Furthermore, we envisage that such data-driven metrics can be adapted to score and compare a variety of FFs, not restricted to lipids, and that this will become particularly useful for analyzing molecular models of complex interacting (e.g. self-assembling) systems.

Methods

Descriptor of Atomic Environments (SOAP)

The Smooth Overlap of Atomic Position (SOAP)^{35,36} aims to accurately reproduce many-body densities of every *site*¹ in a system of interest. In particular, SOAP is a density-based method to encode molecular environments coming from a simulation into a roto-translational invariant representation given by a vector, commonly called “SOAP power spectrum”.³⁵ This is done by expanding the local density $\rho_i(\mathbf{r})$ (defined in a neighborhood of every center by a spatial cutoff r_{cut}) projecting it onto a basis of orthogonal radial functions $g_n(r)$ and spherical harmonics $Y_{lm}(\theta, \phi)$, which for the i -th site can be expressed as

$$\rho_i(\mathbf{r}) = \sum_{j \in r_{cut}} \sum_{nlm} c_{nlm}^j g_n(r) Y_{lm}(\theta, \phi), \quad (1)$$

where the j index runs over all the sites in the cutoff. In practice, Eq. 1 can be analytically solved giving the so called SOAP power spectra vector

$$\mathbf{p}(\mathbf{r})_{nn'l} = \pi \sqrt{\frac{8}{2l+1}} \sum_{m=-l}^l c_{nlm}^*(\mathbf{r}) c_{n'l m}(\mathbf{r}), \quad (2)$$

which encodes all the information of the atomic environment. Eq. 2 represents also the computational output obtained from the SOAP calculation using the DSCRIBE⁶¹ package. Additional details and explanation can be found on the Supplemental Information, but for extensive mathematical derivations, we refer the reader to the original article on the SOAP method.³⁵

A similarity measure between two environments centered in two sites can be formally defined by building a linear kernel of their density representations. Such kernel can be

¹The site is intended as the center of application of the SOAP descriptor, which might be a single atom or pseudo-atom, a set of atoms or a fictitious point, like the center of mass of a set of atoms, depending on the system.

analytically computed, and it can be reduced to the dot product of the two sites' SOAP power spectra.³⁵

$$K^{SOAP}(i, j) = \rho_i(\mathbf{r}) \cdot \rho_j(\mathbf{r}) \propto \mathbf{p}_i \cdot \mathbf{p}_j. \quad (3)$$

Eq. 3 can be interpreted as a measure of how much the two environments are superimposed to each other (i.e., how similar they are). The value of K^{SOAP} goes from 0 for completely different to 1 for matching environments.

The power spectra and Eq. 3 can be further exploited to define a straightforward similarity metrics between two sites via the definition of a SOAP distance

$$d^{SOAP}(i, j) = \sqrt{2 - 2 \cdot K^{SOAP}(i, j)} \propto \sqrt{2 - 2\mathbf{p}_i \cdot \mathbf{p}_j}, \quad (4)$$

where $\mathbf{p}_i$ is the i -th center's power spectrum. Both kernel and distance representation give a bounded measure of how similar two sites are (i.e. how their local densities are orthogonal).

SOAP calculation and parametrization

For all systems, the SOAP descriptors were calculated in periodic boundary condition along the xy dimensions using $n_{max} = 8$ radial basis function and $L_{max} = 8$ maximum degree of spherical harmonics (a gaussian density smoothing $\sigma = 0.1$ nm was used). The SOAP descriptors were centered on the phosphate group atoms, including also the polar organic group atoms of the hydrophilic head (i.e., choline) and the two tails terminal bead in the environment (or, in the case of AA and UA, the atoms corresponding to the last bead in CG representation), thus considering also the environment of the lipid tails inside the membrane. Besides the parameters for the harmonic and radial expansions, the last arbitrary free parameter to select for SOAP calculation was the cut-off distance. We selected a value of 3 nm on the basis of two considerations: (i) the cutoff needs to be large enough to include all particles intended to be characterized by SOAP descriptors; the most demanding criteria is for the 4 beads system for which the cut off needs to be smaller than the width of the

membrane to exclude the opposite surface beads, thus smaller than about 3.7 nm, but large enough to include the tail beads, located in the mid-part of the membrane, thus larger than about 2 nm. (ii) ideally, the descriptors should provide an optimal cut-off range beyond which any further refinement does not add information (or, in other word, the presence of a limit persistence length). We observed that the effective rank of the distance kernel matrix as a function of the SOAP cut off distance for the set of system plateaus at about 3 nm (Figure SI2), consistently with the physical-inspired consideration. All the SOAP calculations were done using the DScibe⁶¹ python package. The complete set of input files and scripts used to analyze the trajectories is available on GitHub at <https://github.com/GMPavanLab/lipids-comparison>.

SOAP comparison

SOAP descriptors are rotational invariant embedding of the environment surrounding a given particle. In order to compare the conformational space originated by the different force fields via MD simulations, one can consider to: (i) use a similarity kernel on the average SOAP descriptors calculated over the whole trajectory, as we explained in the sections above; or (ii) estimate the probability densities resulting from the ensemble of the environments' SOAP power spectra over a reduced representation, and calculate a distributional distance over such densities. To obtain a comparison of type (ii), we started estimating the intrinsic dimension of the SOAP spectra dataset, via the TwoNN algorithm⁶² and with FCI algorithm⁶³ and it was estimated to be larger than 20, for the chosen 3nm cut-off. To make the gridding computationally feasible, we first employed PCA to reduce the dimensionality of the dataset, and used the Pak algorithm⁶⁴ to obtain a uniform grid (details in the SI). As per the available python implementation, the out of sample estimate of the density was carried out by fitting an additional Nearest Neighbor search tree and averaging (via distance weighting) the density values over the first three neighbors. This procedure was considered adequate, given that the extrapolation was carried out over a refined and filtered grid to reduce the

risk of extrapolating over sparsely sampled points of the phase space. Once the probability density was calculated for all system over the grid, the Jensen-Shannon divergence was calculated for comparison with SOAP distance (Figure S1 in the SI).

Probabilistic Analysis of Molecular Motifs (PAMM)

Probabilistic Analysis of Molecular Motifs (PAMM)⁴² is a density-based clustering technique developed with the core idea of partitioning the Probability Distribution Function (PDF) of data sampled from molecular simulations to identify structural motifs. The input required is a set of vectors that completely characterize the molecular environment of the system under analysis (in our case, SOAP power spectra). From these high-dimensional data, PAMM workflow starts from an iterative Kernel Density Estimation (KDE) of the data. Once a stable estimation of the PDF is obtained, a density based clustering is performed, identifying the different local maxima in the PDF, that define cluster centroids. Each cluster is thus fit using a Gaussian Mixture Model (GMM) that ultimately gives the so called Probabilistic Motifs Identifiers (PMIs)⁴² that translates in the system specific fingerprint. The clustering analyses for our systems were performed using the original PAMM code⁴² (available online at <https://github.com/cosmo-epfl/pamm>) as a baseline, along with a tailored Python3 code wrapper to handle the different steps in the analysis workflow and the data post-processing.

Motifs analysis

All the DPPC lipid bilayers were simulated for a total of 1 μs of simulation time at the selected temperatures. SOAP local descriptors were calculated from the production trajectories considering one snapshot every 10 ns (for a total number of 101 frames). We initially merged all the data coming from the same FF simulations, performing a dimensionality reduction performed with a linear PCA, limiting the estimation to the first 5 eigenvectors. In this way, all the temperatures are in the same parameter space and can be directly compared. We thus performed the PAMM clustering on this low-dimensional space and we obtain the clus-

ters separation, and interconversion matrices shown in Figure 4. The complete set of input files and scripts used for the PAMM unsupervised clustering and motifs analyses of the MD trajectories is available on GitHub at <https://github.com/GMPavanLab/lipids-comparison>.

Acknowledgement

The authors thank Exequiel E. Barrera and Loris Di Cairano for useful discussion. G.M.P. acknowledges the funding received by the Swiss National Science Foundation (SNSF grants IZLIZ2_183336) and by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 818776 - DYNAPOL). This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s934.

Supporting Information Available

The complete set of input files and scripts used to analyze the trajectories is available on GitHub at <https://github.com/GMPavanLab/lipids-comparison>.

All the simulation parameters, some considerations on SOAP analysis and some supplementary figures are in the Supplemental Information file.

References

- (1) Van Meer, G.; Voelker, D. R.; Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 112–124.
- (2) Mathai, J. C.; Tristram-Nagle, S.; Nagle, J. F.; Zeidel, M. L. Structural determinants of water permeability through the lipid membrane. *J. Gen. Physiol.* **2008**, *131*, 69–76.
- (3) Sunshine, H.; Iruela-Arispe, M. L. Membrane lipids and cell signaling. *Curr. Opin. Lipidol.* **2017**, *28*, 408.
- (4) Corradi, V.; Sejdiu, B. I.; Mesa-Galloso, H.; Abdizadeh, H.; Noskov, S. Y.; Marrink, S. J.; Tieleman, D. P. Emerging diversity in lipid–protein interactions. *Chem. Rev.* **2019**, *119*, 5775–5848.
- (5) Seelig, A.; Seelig, J. Dynamic structure of fatty acyl chains in a phospholipid bilayer measured by deuterium magnetic resonance. *Biochemistry* **1974**, *13*, 4839–4845.
- (6) Ulrich, A. S.; Sami, M.; Watts, A. Hydration of DOPC bilayers by differential scanning calorimetry. *Biochim. Biophys. Acta, Biomembr.* **1994**, *1191*, 225–230.
- (7) Kučerka, N.; Nagle, J. F.; Sachs, J. N.; Feller, S. E.; Pencer, J.; Jackson, A.; Katsaras, J. Lipid bilayer structure determined by the simultaneous analysis of neutron and X-ray scattering data. *Biophys. J.* **2008**, *95*, 2356–2367.
- (8) Zhang, R.; Tristram-Nagle, S.; Sun, W.; Headrick, R.; Irving, T.; Suter, R. M.; Nagle, J. F. Small-angle x-ray scattering from lipid bilayers is well described by modified Caillé theory but not by paracrystalline theory. *Biophys. J.* **1996**, *70*, 349–357.
- (9) Ingólfsson, H. I.; Arnarez, C.; Periole, X.; Marrink, S. J. Computational ‘microscopy’ of cellular membranes. *J. Cell Sci.* **2016**, *129*, 257–268.

- (10) Marrink, S. J.; Corradi, V.; Souza, P. C.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. Computational modeling of realistic cell membranes. *Chem. Rev.* **2019**, *119*, 6184–6226.
- (11) Tieleman, D. P.; MacCallum, J. L.; Ash, W. L.; Kandt, C.; Xu, Z.; Monticelli, L. Membrane protein simulations with a united-atom lipid and all-atom protein model: lipid–protein interactions, side chain transfer free energies and model proteins. *J. Phys.: Condens. Matter* **2006**, *18*, S1221.
- (12) Bennun, S. V.; Hoopes, M. I.; Xing, C.; Faller, R. Coarse-grained modeling of lipids. *Chem. Phys. Lipids* **2009**, *159*, 59–66.
- (13) Srivastava, A.; Voth, G. A. Hybrid approach for highly coarse-grained lipid bilayer models. *J. Chem. Theory Comput.* **2013**, *9*, 750–765.
- (14) Siu, S. W.; Vácha, R.; Jungwirth, P.; Böckmann, R. A. Biomolecular simulations of membranes: physical properties from different force fields. *J. Chem. Phys.* **2008**, *128*, 03B620.
- (15) Pluhackova, K.; Kirsch, S. A.; Han, J.; Sun, L.; Jiang, Z.; Unruh, T.; Böckmann, R. A. A critical comparison of biomembrane force fields: structure and dynamics of model DMPC, POPC, and POPE bilayers. *J. Phys. Chem. B* **2016**, *120*, 3888–3903.
- (16) Botan, A.; Favela-Rosales, F.; Fuchs, P. F.; Javanainen, M.; Kanduč, M.; Kulig, W.; Lamberg, A.; Loison, C.; Lyubartsev, A.; Miettinen, M. S. et al. Toward atomistic resolution structure of phosphatidylcholine headgroup and glycerol backbone at different ambient conditions. *J. Phys. Chem. B* **2015**, *119*, 15075–15088.
- (17) Ollila, O. S.; Pabst, G. Atomistic resolution structure and dynamics of lipid bilayers in simulations and experiments. *Biochim. Biophys. Acta, Biomembr.* **2016**, *1858*, 2512–2528.

- (18) Jämbeck, J. P.; Lyubartsev, A. P. Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids. *J. Phys. Chem. B* **2012**, *116*, 3164–3179.
- (19) Jämbeck, J. P.; Lyubartsev, A. P. An extension and further validation of an all-atomistic force field for biological membranes. *J. Chem. Theory Comput.* **2012**, *8*, 2938–2948.
- (20) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O’Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell Jr, A. D.; Pastor, R. W. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **2010**, *114*, 7830–7843.
- (21) Gould, I.; Skjervik, A.; Dickson, C.; Madej, B.; Walker, R. Lipid17: A Comprehensive AMBER Force Field for the Simulation of Zwitterionic and Anionic Lipids. *Manuscript in preparation* **2018**,
- (22) Berger, O.; Edholm, O.; Jähnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* **1997**, *72*, 2002–2013.
- (23) Chiu, S.-W.; Pandit, S. A.; Scott, H.; Jakobsson, E. An improved united atom force field for simulation of mixed lipid bilayers. *J. Phys. Chem. B* **2009**, *113*, 2748–2763.
- (24) Piggot, T. J.; Piñeiro, A.; Khalid, S. Molecular dynamics simulations of phosphatidylcholine membranes: a comparative force field study. *J. Chem. Theory Comput.* **2012**, *8*, 4593–4609.
- (25) Marrink, S. J.; De Vries, A. H.; Mark, A. E. Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* **2004**, *108*, 750–760.
- (26) 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* **2007**, *111*, 7812–7824.
- (27) Barrera, E. E.; Machado, M. R.; Pantano, S. Fat SIRAH: Coarse-Grained Phospholipids To Explore Membrane–Protein Dynamics. *J. Chem. Theory Comput.* **2019**, *15*, 5674–5688.
- (28) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **2010**, *6*, e1000810.
- (29) de Jong, D. H.; Singh, G.; Bennett, W. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, 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.
- (30) Khan, H. M.; Souza, P. C.; Thallmair, S.; Barnoud, J.; De Vries, A. H.; Marrink, S. J.; Reuter, N. Capturing Choline–Aromatics Cation- π Interactions in the MARTINI Force Field. *J. Chem. Theory Comput.* **2020**, *16*, 2550–2560.
- (31) Arnarez, C.; Uusitalo, J. J.; Masman, M. F.; Ingólfsson, H. I.; De Jong, D. H.; Melo, M. N.; Periole, X.; De Vries, A. H.; Marrink, S. J. Dry Martini, a coarse-grained force field for lipid membrane simulations with implicit solvent. *J. Chem. Theory Comput.* **2015**, *11*, 260–275.
- (32) Ceriotti, M.; Tribello, G. A.; Parrinello, M. Simplifying the representation of complex free-energy landscapes using sketch-map. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13023–13028.
- (33) Artrith, N.; Urban, A.; Ceder, G. Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species. *Phys. Rev. B* **2017**, *96*, 014112.
- (34) Engel, E. A.; Anelli, A.; Ceriotti, M.; Pickard, C. J.; Needs, R. J. Mapping uncharted territory in ice from zeolite networks to ice structures. *Nat. Comm.* **2018**, *9*, 1–7.

- (35) Bartók, A. P.; Kondor, R.; Csányi, G. On representing chemical environments. *Phys. Rev. B* **2013**, *87*, 184115.
- (36) De, S.; Bartók, A. P.; Csányi, G.; Ceriotti, M. Comparing molecules and solids across structural and alchemical space. *Phys. Chem. Chem. Phys.* **2016**, *18*, 13754–13769.
- (37) Monserrat, B.; Brandenburg, J. G.; Engel, E. A.; Cheng, B. Liquid water contains the building blocks of diverse ice phases. *Nat. Comm.* **2020**, *11*, 1–8.
- (38) Cheng, B.; Griffiths, R.-R.; Wengert, S.; Kunkel, C.; Stenczel, T.; Zhu, B.; Deringer, V. L.; Bernstein, N.; Margraf, J. T.; Reuter, K. et al. Mapping materials and molecules. *Acc. Chem. Res.* **2020**, *53*, 1981–1991.
- (39) Gasparotto, P.; Bochicchio, D.; Ceriotti, M.; Pavan, G. M. Identifying and tracking defects in dynamic supramolecular polymers. *J. Phys. Chem. B* **2020**, *124*, 589–599.
- (40) de Marco, A. L.; Bochicchio, D.; Gardin, A.; Doni, G.; Pavan, G. M. Controlling Exchange Pathways in Dynamic Supramolecular Polymers by Controlling Defects. ChemRxiv, 2021, DOI: 10.26434/chemrxiv.13655864.v2.
- (41) Lin, J. Divergence measures based on the Shannon entropy. *IEEE Trans. Inf. Theory* **1991**, *37*, 145–151.
- (42) Gasparotto, P.; Meißner, R. H.; Ceriotti, M. Recognizing Local and Global Structural Motifs at the Atomic Scale. *J. Chem. Theory Comput.* **2018**, *14*, 486–498, PMID: 29298385.
- (43) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y. et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **2016**, *12*, 405–413.

- (44) Hsu, P.-C.; Bruininks, B. M.; Jefferies, D.; Cesar Telles de Souza, P.; Lee, J.; Patel, D. S.; Marrink, S. J.; Qi, Y.; Khalid, S.; Im, W. CHARMM-GUI Martini Maker for modeling and simulation of complex bacterial membranes with lipopolysaccharides. *J. Comput. Chem.* **2017**, *38*, 2354–2363.
- (45) Ferreira, T. M.; Coreta-Gomes, F.; Ollila, O. S.; Moreno, M. J.; Vaz, W. L.; Topgaard, D. Cholesterol and POPC segmental order parameters in lipid membranes: solid state ^1H – ^{13}C NMR and MD simulation studies. *Phys. Chem. Chem. Phys.* **2013**, *15*, 1976–1989.
- (46) Samuli, O. O. H.; Tiago, F.; Daniel, T. MD simulation trajectory and related files for POPC bilayer (Berger model delivered by Tieleman, Gromacs 4.5). 2014; <https://doi.org/10.5281/zenodo.13279>.
- (47) Domański, J.; Stansfeld, P. J.; Sansom, M. S.; Beckstein, O. Lipidbook: a public repository for force-field parameters used in membrane simulations. *J. Membr. Biol.* **2010**, *236*, 255–258.
- (48) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19–25.
- (49) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604–613.
- (50) Bonomi, M.; Bussi, G.; Camilloni, C.; Tribello, G. A.; Banáš, P.; Barducci, A.; Bernetti, M.; Bolhuis, P. G.; Bottaro, S.; Branduardi, D. et al. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670–673.
- (51) Poger, D.; Caron, B.; Mark, A. E. Validating lipid force fields against experimental data: Progress, challenges and perspectives. *Biochim. Biophys. Acta, Biomembr.* **2016**, *1858*, 1556–1565.

- (52) Kučerka, N.; Tristram-Nagle, S.; Nagle, J. F. Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains. *J. Membr. Biol.* **2006**, *208*, 193–202.
- (53) 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, Biomembr.* **2011**, *1808*, 2761–2771.
- (54) 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.* **2021**, *17*, 1170–1180.
- (55) Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* **1964**, *29*, 1–27.
- (56) Bochicchio, D.; Salvalaglio, M.; Pavan, G. M. Into the dynamics of a supramolecular polymer at submolecular resolution. *Nat. Comm.* **2017**, *8*, 147.
- (57) Zhao, J.; Wu, J.; Shao, H.; Kong, F.; Jain, N.; Hunt, G.; Feigenson, G. Phase studies of model biomembranes: macroscopic coexistence of $L\alpha + L\beta$, with light-induced coexistence of $L\alpha + L_o$ phases. *Biochim. Biophys. Acta, Biomembr.* **2007**, *1768*, 2777–2786.
- (58) Biltonen, R. L.; Lichtenberg, D. The use of differential scanning calorimetry as a tool to characterize liposome preparations. *Chem. Phys. Lipids* **1993**, *64*, 129–142.
- (59) Marrink, S. J.; Risselada, J.; Mark, A. E. Simulation of gel phase formation and melting in lipid bilayers using a coarse grained model. *Chem. Phys. Lipids* **2005**, *135*, 223–244.
- (60) Stelter, D.; Keyes, T. Enhanced sampling of phase transitions in coarse-grained lipid bilayers. *J. Phys. Chem. B* **2017**, *121*, 5770–5780.
- (61) Himanen, L.; Jäger, M. O. J.; Morooka, E. V.; Federici Canova, F.; Ranawat, Y. S.; Gao, D. Z.; Rinke, P.; Foster, A. S. DDescribe: Library of descriptors for machine learning in materials science. *Comput. Phys. Commun.* **2020**, *247*, 106949.

- (62) Facco, E.; D’Errico, M.; Rodriguez, A.; Laio, A. Estimating the intrinsic dimension of datasets by a minimal neighborhood information. *Sci. Rep.* **2017**, *7*, 1–8.
- (63) Erba, V.; Gherardi, M.; Rotondo, P. Intrinsic dimension estimation for locally under-sampled data. *Sci. Rep.* **2019**, *9*, 1–9.
- (64) Rodriguez, A.; D’Errico, M.; Facco, E.; Laio, A. Computing the Free Energy without Collective Variables. *J. Chem. Theory Comput.* **2018**, *14*, 1206–1215.

Supporting Information for:
A Data-Driven Dimensionality Reduction
Approach to Compare and Classify Lipid Force
Fields

Riccardo Capelli^{1,*}, Andrea Gardin¹, Charly Empereur-mot²,
Giovanni Doni², and Giovanni M. Pavan^{1,2,*}

¹Department of Applied Science and Technology, Politecnico di
Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy

²Department of Innovative Technologies, University of Applied
Sciences and Arts of Southern Switzerland, Polo Universitario
Lugano - Campus Est, Via la Santa 1, CH-6962 Lugano -
Viganello, Switzerland

*riccardo.capelli@polito.it; giovanni.pavan@polito.it

S1 Lipid Bilayer Observables

Here we briefly summarize the observables we have evaluated. To obtain those observables from our MD trajectories, we employed PLUMED,^{1,2} MDAnalysis³ and standard GROMACS⁴ tools.

S1.1 Area Per Lipid (A_L)

Area per Lipid A_L is defined as

$$A_L = \frac{B_x B_y}{N_L}$$

where B_x and B_y are the width of the simulation box in the x and the y axis, respectively, and N_L is the number of lipid molecules in a single layer.

S1.2 Membrane Thickness (D_{HH})

Membrane thickness D_{HH} is given by the average distance of the phosphate groups projected along the direction normal to the plane define by the bilayer

(z axis in our case), namely:

$$D_{HH} = \langle z^{\text{up}} \rangle - \langle z_j^{\text{down}} \rangle$$

where $\langle z^{\text{up}} \rangle$ is the average of the z -position of the phosphate group in the upper layer, and $\langle z^{\text{down}} \rangle$ is the average of the z -position of the phosphate group in the lower layer.

S1.3 Membrane Compressibility (K_A)

Membrane compressibility K_A is defined as

$$K_A = A \left(\frac{\partial \gamma}{\partial A} \right)_T = \frac{A_L k_B T}{N_L \sigma_A^2}$$

where A is the total bilayer area, γ is the surface tension, k_B is the Boltzmann constant, T is the temperature, N_L is the number of lipids in a single layer, A_L is the area per lipid, and σ_A^2 is the mean square fluctuation of A_L .

S1.4 Lateral Diffusion Coefficient ($\mathcal{D}_l$)

Lateral Diffusion Coefficient $\mathcal{D}_l$ can be computed exploiting the Einstein relation for a diffusive particle in a 2-dimensional space

$$\mathcal{D}_l = \frac{1}{4} \lim_{t \rightarrow \infty} \frac{\langle [\Delta \vec{r}]^2 \rangle}{t}$$

where $\langle [\Delta \vec{r}]^2 \rangle$ is the mean square displacement (MSD) for all the lipids in the membrane. Practically speaking, for long lag times t , it is possible to obtain $\mathcal{D}_l$ as the slope of $\langle [\Delta \vec{r}]^2 \rangle$:

$$\langle [\Delta \vec{r}]^2 \rangle = 4\mathcal{D}_l t$$

The number obtained from a MD simulation using this relation is often called $\mathcal{D}_l^{pbc}$. This is due to the fact that the presence of periodic boundary condition cause a systematic deviation of the diffusion coefficient values. Following hydrodynamic theory, it is possible to obtain an analytical correction.⁵⁻⁷ In this work, we computed and showed $\mathcal{D}_l^{pbc}$ values only to prove the variability of this observable in different force field parametrization. A corrected (and more precise) evaluation of $\mathcal{D}_l$ is beyond the scope of this work.

S2 Smooth Overlap of Atomic Position

The Smooth Overlap of Atomic Position (SOAP)⁸ is a fairly recent many-body atomic descriptor, that aims to accurately reproduce atomic or in general particle densities of a given system of interest. The description obtained by SOAP, commonly called “atomic-environment” is usually encoded in a high dimensional vector, that stores all the information regarding a specific center of application of the descriptor. Others known examples of atomic descriptors are:

Atom-centered Symmetry Functions (ACSFs),⁹ Coulomb Matrix (CM),¹⁰ sine matrix,¹¹ many-body tensor representation (MBTR).¹² All the notations and terminology about SOAP were first introduced by Bartok *et al.*,⁸ where the authors extensively described and validated all the mathematical and physical properties of such framework. In this section we will include some basic notations and expressions to help the reader get familiar with such formalism, but we will not enter in too much detail on the framework, since it deviates from the scope of our article¹.

Let a generic system of interest be composed of a set $\mathcal{G}$ of $I = (1, \dots, i, \dots, N)$ atomic sites², given a cutoff sphere $S(r_{cut})$ the i th-particle density is formally defined as

$$\rho^{(i)}(\mathbf{r}) = \sum_a \rho_a^{(i)}(\mathbf{r}) = \sum_a \delta(\mathbf{r} - \mathbf{r}_a), \quad (\text{S1})$$

where the index a runs over a subset $\mathcal{A} \subseteq \mathcal{G}$ of sites included in the spherical cutoff $S(r_{cut})$. In the SOAP formalism⁸ the full representation of the density function (Eq. S1) is expressed combining a set of radial basis $\{g_n\}$ and spherical harmonics $Y_{lm}(\theta, \phi)$ functions

$$\rho^{(i)}(\mathbf{r}) \simeq \tilde{\rho}^{(i)}(\mathbf{r}) = \sum_a \sum_{nlm} c_{nlm}^a g_n(r) Y_{lm}(\theta, \phi), \quad (\text{S2})$$

where c_{nlm}^a are the spherical harmonics and radial functions' coefficients, that cycles over the indices n (the index related to the radial expansion), l and m (the indices related to the spherical harmonics expansion). $\tilde{\rho}^{(i)}(\mathbf{r})$ is also commonly referred as Gaussian smoothed density function. The latter allows the construction of a description of the local density of the atomic site, that in general approximates well the estimate of $\rho^{(i)}(\mathbf{r})$ (Eq. S1) and retains its permutation invariant property (from now on, we will refer to $\tilde{\rho}$ as simply ρ to simplify our notation).

Given this approximation, in ref.⁸ it is shown that to keep a local density that results rotationally invariant, one can consider the so-called “power spectrum” of Eq. S2, which expression is

$$\mathbf{p}(\mathbf{r})_{nn'l} = \pi \sqrt{\frac{8}{2l+1}} \sum_{m=-l}^l c_{nlm}^*(\mathbf{r}) c_{n'lm}(\mathbf{r}), \quad (\text{S3})$$

with the coefficients $c_{n'lm}$ are defined as the inner product

$$c_{n'lm} = \langle g_n Y_{lm} | \rho \rangle. \quad (\text{S4})$$

The vector defined in Eq. S3 is formally referred as the partial power spectrum of the Gaussian smeared atomic density, which defines the SOAP vector and it represent also the computational output obtained from the SOAP

¹For the interested reader, we would like to suggest the reading of ref.^{8,13}

²We use the general term sites to indicate both real and pseudo atoms used as centers for the SOAP representation.

calculation using the DDescribe¹⁴ package. A schematic representation of the computational workflow is depicted in Fig. S2. All the SOAP calculations were done using the DDescribe¹⁴ python package. The complete set of input files and scripts used to analyze the trajectories is available on GitHub at <https://github.com/GMPavanLab/lipids-comparison>.

S2.1 SOAP similarity measure

Given the mathematical description of SOAP, it appears natural to consider the possibility to compare different atomic environments. This task can be done through the definition of a (dis)similarity measure (*i.e.*, a metric) although others strategies have been already explored.¹³

In Ref.⁸ the authors laid the basis of a straightforward (dis)similarity measure between two different environments. Operatively, this metrics is defined as inner product of two densities,⁸ namely

$$I(i, j) = \int \rho^{(i)}(\mathbf{r}) \rho^{(j)}(\mathbf{r}) d\mathbf{r}. \quad (\text{S5})$$

Eq. S5 inherits from the definition of $\rho(\mathbf{r})$ its permutation invariance. Rotation invariance of this metrics is taken into account by integrating Eq. S5 over all possible rotations, leading to the definition of a SOAP similarity linear kernel

$$\begin{aligned} \tilde{K}(i, j) &= \int |I(i, \hat{R}j)|^\zeta d\hat{R} \\ &= \int d\hat{R} \left| \int \rho^{(i)}(\mathbf{r}) \rho^{(j)}(\hat{R}\mathbf{r}) d\mathbf{r} \right|^\zeta, \end{aligned} \quad (\text{S6})$$

where $\hat{R}$ represent a generic rotation, the tilde superscript means that the kernel is not unit-normalized and ζ is an integer that needs to be $\zeta \geq 2$ in order to maintain the angular information.⁸

In all our calculations the ζ -constant was kept equal to two. Recalling the definition of SOAP power spectrum (Eq. S3), after some algebraic manipulation³, we can rewrite Eq. S6 as

$$\tilde{K}(i, j) = \tilde{\mathbf{p}}_i \cdot \tilde{\mathbf{p}}_j, \quad (\text{S7})$$

where $\tilde{\mathbf{p}}_i$ represents the non-normalized SOAP power spectrum vector of the i -th environment.⁸ To make all the kernels commensurable, we decided to consider the unit-normalized kernel, defined as

$$K(i, j) = \frac{\tilde{K}(i, j)}{\sqrt{\tilde{K}(i, i) \tilde{K}(j, j)}} = \frac{\tilde{\mathbf{p}}_i \cdot \tilde{\mathbf{p}}_j}{\sqrt{(\tilde{\mathbf{p}}_i \cdot \tilde{\mathbf{p}}_i)(\tilde{\mathbf{p}}_j \cdot \tilde{\mathbf{p}}_j)}}. \quad (\text{S8})$$

$K(i, j)$ gives a similarity measure that can vary from 0 (the environments are not superimposed) to 1 (the environments are completely superimposed).

³Complete derivation is available in ref.⁸

Figure S1: Schematic representation of the SOAP analysis. a) centers selection of a sample Martini 2.2 lipid and the definition of the single lipid Gaussian smoothed density profile. From left to right: the asterisks identify the 4 sites (3 kinds of bead) considered by SOAP; concentric circles of different colors identify the application of the SOAP density profiles on every center. b) at the assembly level, the SOAP lipid environment is completed by considering all the other lipids contribution inside a specified r_{cut} centered in the lipid of choice (white asterisk). c) computation through the DScibe library¹⁴ of the SOAP power spectra which define the complete lipid environment, up to the chosen cutoff. The target lipid (identified by the asterisk) representation is affected by all the others in its surroundings, defining its local environment. This schematic procedure is repeated for all the lipid in a target membrane to obtain its complete local characterization (lipid-wise).

S2.2 Considerations on the SOAP metrics

The definition of the SOAP similarity kernel in terms of a dot product (Eq. S7) leads naturally to a definition of a similarity metric. Let $\mathbf{a} = (a_1, a_2, \dots, a_D)$ and $\mathbf{b} = (b_1, b_2, \dots, b_D)$ be two points in Euclidean D-space, their distance is given by

$$d(\mathbf{a}, \mathbf{b}) = d(\mathbf{b}, \mathbf{a}) = \sqrt{\sum_d^D (b_k - a_k)^2}. \quad (\text{S9})$$

The position of a point in a Euclidean D-space is tied to the knowledge of Euclidean vector seen as a directed line segment from the space frame origin (tail) to a point in that space (tip). Its magnitude is simply its norm, $\|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$, which can be seen as the distance between the vector tail and tip. If the relationship between $\mathbf{a}$ and $\mathbf{b}$ involve a direction, this relation is expressed simply as $(\mathbf{a} - \mathbf{b}) = (a_1 - b_1, \dots, a_D - b_D)$, which is called the displacement vector from $\mathbf{b}$ to $\mathbf{a}$. The Euclidean distance between $\mathbf{a}$ and $\mathbf{b}$ then, is just the Euclidean length of this displacement vector

$$\begin{aligned} d(\mathbf{b}, \mathbf{a}) &= \|\mathbf{b} - \mathbf{a}\| = \sqrt{(\mathbf{b} - \mathbf{a}) \cdot (\mathbf{b} - \mathbf{a})} = \\ &= \sqrt{\|\mathbf{b}\|^2 + \|\mathbf{a}\|^2 - 2\mathbf{b} \cdot \mathbf{a}} \end{aligned} \quad (\text{S10})$$

At this point combining Eq. S10 with the definition of SOAP kernel (Eq. S7) we can measure the similarity distance between two environments as

$$\begin{aligned} d^{SOAP}(i, j) &= \sqrt{K(i, i) + K(j, j) - 2K(i, j)} = \sqrt{2 - 2K(i, j)} \\ &= \sqrt{2 - 2\mathbf{p}_i \cdot \mathbf{p}_j} \end{aligned} \quad (\text{S11})$$

Recently, the definition of this metric and some derivations from it proven to be a useful tool in evaluating the similarity of crystalline, disordered and molecular compounds.¹³

S2.3 Reducing the dimensionality of SOAP descriptors

Irrespective of the cutoff value, the dimensionality of the SOAP spectra calculated using both the TwoNN algorithm¹⁵ and the FCI algorithm¹⁶ is greater than 20. It is clear that the size of the raw SOAP embedding for the 4 bead reduced representation is highly redundant. In fact, 80% of the variance is accounted for by just 5 pca components as shown in.S1. PCA was used as dimensionality reduction technique throughout the work, for its its efficiency and low computational cost.

Table S1: SOAP dimensionality reduction				
Cut-off	SOAP dim	ID	5 pca	% variance
20	2700	23		59.1
30	2700	24		79.1

S2.4 Jensen-Shannon divergence

To validate the average SOAP distance metric, we computed the Jansen-Shannon divergence over the probability densities samples by MD for all systems.

In order to do so we proceed as follows. Firstly, the SOAP spectra were projected to the first 5 dimensional component calculated using a dataset composed of samples of SOAP spectra for each systems. The probability density for each system was estimated via the non parametric Pak algorithm⁷ on 20 bootstrapped samples in the reduced representation.

The Jansen-Shannon divergence was then calculated between the probabilities densities extrapolated over a much finer grid for all FF pairs. The extrapolation of out of sample data is not dealt directly by the available Pak implementation and was carried out by applying a distance waiting criteria on neighbouring; additionally, the lack of a asymptotic decay intrinsic of the non parametric nature of the estimator, was dealt with by applying a smoothing term.

To avoid memory issue resulting from the number of grid points required (about n^d , for a discretization into n intervals) and avoid estimating probabilities over scarcely sampled areas of the domain, the grid was generated in a lazy manner and was progressively filtered via a Multilabel Perceptron classifier, tuned to strip out grid points likely to belong to the discretization (hence outside the range of values generated by the various ensembles) and not to any of the force fields systems.

The use of alternative to this strategy for the determination of the probability densities making use of Variational Autonecoders (thus providing the possibility of calculating divergences out-of-the-box) have been considered and deemed beyond the score of the present works.

S3 PAMM clustering

The analysis of molecular motifs was carried out following the same workflow for the two data sets of lipid membranes. First, the collection of all the SOAP vectors for all the five temperatures of specific force field was reduced to a sub-set using a *farthest point* sampling method. We used 2000 representative points, sampled non uniformly in this way for the grid definition. The KDE was then carried out to select the underlying modes. We used as parameters for the clustering `input-dim = 5`, `fspread = 0.3`, `quick-shift = 1`, `merger = 0.001` and a total of 73 bootstrap runs for the statistical calculations. The complete set of input files and scripts used for the PAMM unsupervised clustering and motifs analyses of the MD trajectories is available onGitHub at <https://github.com/GMPavanLab/lipids-comparison>.

S4 Simulation details

S4.1 All-atom Force Fields

To prepare the initial configurations for Charmm36 and Slipids force fields we used the web tool CHARMM-GUI,¹⁷ putting a 50 Å-high layer of TIP3P¹⁸ water for both sides of the membrane without counterions. For AMBER LIPID17, we took one bilayer configuration without water from Charmm36 system preparation and using tleap from AmberTools20,¹⁹ we solvated it and converted the model to GROMACS format using Parmed.²⁰ We thus followed the minimization and equilibration protocol of CHARMM-GUI, namely:

- 5000 steps of Steepest descent minimization;
- 2.5e5 steps (timestep of 0.5 fs, 125 ps) of MD and position (1000 kJ/mol/nm²) and dihedral restraints (1000 kJ/mol/rad²) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of 4.5×10^{-5} bar⁻¹) and thermostat²¹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 1.25e5 steps (timestep of 1 fs, 125 ps) of MD and position (1000 kJ/mol/nm²) and dihedral restraints (400 kJ/mol/rad²) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of 4.5×10^{-5} bar⁻¹) and thermostat²¹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 1.25e5 steps (timestep of 1 fs, 125 ps) of MD and position (400 kJ/mol/nm²) and dihedral restraints (200 kJ/mol/rad²) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of 4.5×10^{-5} bar⁻¹) and thermostat²¹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 2.5e5 steps (timestep of 2 fs, 500 ps) of MD and position (200 kJ/mol/nm²) and dihedral restraints (200 kJ/mol/rad²) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of 4.5×10^{-5} bar⁻¹) and thermostat²¹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 2.5e5 steps (timestep of 2 fs, 500 ps) of MD and position (40 kJ/mol/nm²) and dihedral restraints (100 kJ/mol/rad²) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of 4.5×10^{-5} bar⁻¹) and thermostat²¹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 2.5e5 steps (timestep of 2 fs, 500 ps) of MD in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of 4.5×10^{-5} bar⁻¹) and thermostat²¹ (reference temperature of 303 K with a coupling constant of 1 ps);

- 5e5 steps (timestep of 2 fs, 1 ns) in NPT ensemble using a Parrinello-Rahman barostat²² (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$) and Nosé-Hoover thermostat²³ (reference temperature of 303 K with a coupling constant of 1 ps).

All the choice regarding electrostatic long-range calculations are the same as in production runs for every AA FF (see table S2). At the end of this minimization protocol we simulated a 100-ns long equilibration run with the same parameters as the final production run (see table S2).

S4.2 United Atom Force Fields

We obtained the initial conformations from the internet:

- Berger FF from Ref;^{24,25}
- Gromos-CKP from Lipidbook²⁶
- Gromos-43a1-s3 from Lipidbook²⁶

we subsequently solvated the system with a 50 Å-high layer of SPC water²⁷ for both sides of the membrane, and applied the all-atom CHARMM-GUI protocol for minimization and equilibration. After this initial protocol, we performed a 100-ns long equilibration run with the same parameters as the final production run (see table S2).

S4.3 Non-polarizable Martini Force Field

To prepare the initial configurations for Martini 2.2 force field we used the web tool CHARMM-GUI,¹⁷ putting a 50 Å-high layer of Martini water²⁸ water for both sides of the membrane without counterions. Martini 3.0 beta 3.2 has the same mapping of Martini 2.2, thus we used the same initial conformation as a starting point. We followed the minimization and equilibration protocol given by CHARMM-GUI, namely:

- 5000 steps of Steepest descent minimization with a soft-core potential on LJ and Coulomb interaction;
- 5000 steps of Steepest descent minimization;
- 5e5 steps (timestep of 2 fs, 1 ns) of MD and position restraint on head group (200 kJ/mol/nm²) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 2e5 steps (timestep of 5 fs, 1 ns) of MD and position restraint on head group (100 kJ/mol/nm²) in NPT ensemble using a Berendsen barostat

(reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps);

- 1e5 steps (timestep of 10 fs, 1 ns) of MD and position restraint on head group (50 kJ/mol/nm^2) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 5e4 steps (timestep of 15 fs, 750 ps) of MD and position restraint on head group (20 kJ/mol/nm^2) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 5e4 steps (timestep of 20 fs, 1 ns) of MD and position restraint on head group (10 kJ/mol/nm^2) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 1.5e6 steps (timestep of 20 fs, 1 ns) of MD in NPT ensemble using a Parrinello-Rahman barostat²² (reference pressure of 1 bar, a coupling constant of 12 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps).

All the choice regarding electrostatic long-range calculations are the same as in production runs (see table S2). At the end of this minimization protocol we simulated a 100-ns long equilibration run with the same parameters as the final production run (see table S2).

S4.4 Polarizable Martini Force Field

To prepare the initial configurations for Martini 2.2p and Martini 2.3p force fields we used the web tool CHARMM-GUI,¹⁷ putting a 50 Å-high layer of POL water³⁰ water for both sides of the membrane without counterions. We followed the minimization and equilibration protocol given by CHARMM-GUI, namely:

- 5000 steps of Steepest descent minimization with a soft-core potential on LJ and Coulomb interaction;
- 5000 steps of Steepest descent minimization;
- 5e5 steps (timestep of 2 fs, 1 ns) of MD and position restraint on head group (200 kJ/mol/nm^2) in NPT ensemble using a Berendsen barostat

(reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps);

- 2e5 steps (timestep of 5 fs, 1 ns) of MD and position restraint on head group (100 kJ/mol/nm²) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 1e5 steps (timestep of 10 fs, 1 ns) of MD and position restraint on head group (50 kJ/mol/nm²) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 5e4 steps (timestep of 15 fs, 750 ps) of MD and position restraint on head group (20 kJ/mol/nm²) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 5e4 steps (timestep of 20 fs, 1 ns) of MD and position restraint on head group (10 kJ/mol/nm²) in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 5 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps);
- 1.5e6 steps (timestep of 20 fs, 1 ns) of MD in NPT ensemble using a Parrinello-Rahman barostat²² (reference pressure of 1 bar, a coupling constant of 12 ps, and a compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$) and velocity rescale thermostat²⁹ (reference temperature of 303 K with a coupling constant of 1 ps).

All the choice regarding electrostatic long-range calculations are the same as in production runs (see table S2). At the end of this minimization protocol we simulated a 100-ns long equilibration run with the same parameters as the final production run (see table S2).

S4.5 Dry Martini Force Field

To prepare the initial configurations for Dry Martini force field we used the web tool CHARMM-GUI,¹⁷ putting a 50 Å-high empty layer on both sides of the membrane without counterions. We followed the minimization and equilibration protocol given by CHARMM-GUI, namely:

- 5000 steps of Steepest descent minimization with a soft-core potential on LJ and Coulomb interaction;

- 5000 steps of Steepest descent minimization;
- 5e5 steps (timestep of 2 fs, 1 ns) of MD and position restraint on head group (200 kJ/mol/nm²) in NVT ensemble with a langevin thermostat (reference temperature of 303 K with a coupling constant of 4 ps);
- 2e5 steps (timestep of 5 fs, 1 ns) of MD and position restraint on head group (100 kJ/mol/nm²) in NVT ensemble with a langevin thermostat (reference temperature of 303 K with a coupling constant of 4 ps);
- 1e5 steps (timestep of 10 fs, 1 ns) of MD and position restraint on head group (50 kJ/mol/nm²) in NVT ensemble with a langevin thermostat (reference temperature of 303 K with a coupling constant of 4 ps);
- 5e4 steps (timestep of 15 fs, 750 ps) of MD and position restraint on head group (20 kJ/mol/nm²) in NVT ensemble with a langevin thermostat (reference temperature of 303 K with a coupling constant of 4 ps);
- 5e4 steps (timestep of 20 fs, 1 ns) of MD and position restraint on head group (10 kJ/mol/nm²) in NVT ensemble with a langevin thermostat (reference temperature of 303 K with a coupling constant of 4 ps);
- 1.5e6 steps (timestep of 20 fs, 1 ns) of MD in NVT ensemble with a langevin thermostat (reference temperature of 303 K with a coupling constant of 4 ps);

All the choice regarding electrostatic long-range calculations are the same as in production runs (see table S2). At the end of this minimization protocol we simulated a 100-ns long equilibration run with the same parameters as the final production run (see table S2).

S4.6 Sirah Force Field

To prepare the initial configurations for Dry Martini force field we started from the initial POPC bilayer conformation for Charmm36 FF, coarse-graining it following the tutorial on www.sirahff.com. We solvated the system putting a 50 Å-high layer of WAT4 water model³¹ on both sides of the membrane without counterions. We followed the minimization and equilibration protocol given in the Sirah FF bilayer tutorial, namely: To prepare the initial configurations for Dry Martini force field we used the web tool CHARMM-GUI,¹⁷ putting a 50 Å-high empty layer on both sides of the membrane without counterions. We followed the minimization and equilibration protocol given by CHARMM-GUI, namely:

- 100 steps of Steepest descent minimization followed by 4900 steps of conjugate gradient minimization.
- 2.5e4 steps (timestep of 20 fs, 500 ps) of MD in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 8 ps)

and langevin thermostat (reference temperature of 303 K with a coupling constant of 5 ps);

- 2.5e5 steps (timestep of 20 fs, 5 ns) of MD in NPT ensemble using a Berendsen barostat (reference pressure of 1 bar, a coupling constant of 8 ps and langevin thermostat (reference temperature of 303 K with a coupling constant of 5 ps);

All the choice regarding electrostatic long-range calculations are the same as in production runs (see table S2). At the end of this minimization protocol we simulated a 100-ns long equilibration run with the same parameters as the final production run (see table S2).

Table S2: Details in Production Simulation parameters.

Force Field	Time step [ps]	Thermostat τ_T [ps ⁻¹]	Barostat τ_P [ps ⁻¹]	Electrostatic	ES cut-off [nm]	vdW	vdW cut-off [nm]	Water Model		
Slipids Charmm36 LIPID17	0.002	v-rescale ²⁹	1	Parrinello-Rahman ²²	5	PME ³²	1.5	Cut-off/Force-switch	1.4/1.5	TIP3p ¹⁸
	0.002	v-rescale ²⁹	1	Parrinello-Rahman ²²	5	PME ³²	1.2	Cut-off/Force-switch	1.0/1.2	TIP3p ¹⁸
	0.002	v-rescale ²⁹	1	Parrinello-Rahman ²²	5	PME ³²	1.2	Cut-off/Force-switch	1.0/1.2	TIP3p ¹⁸
Berger	0.002	v-rescale ²⁹	0.1	Parrinello-Rahman ²²	8	PME ³²	1.0	Cut-off/Verlet	1.0	SPC ²⁷
Gromos43a1-s3	0.002	v-rescale ²⁹	0.1	Parrinello-Rahman ²²	8	PME ³²	1.0	Cut-off/Verlet	1.0	SPC ²⁷
Gromos-CKP	0.002	v-rescale ²⁹	0.1	Parrinello-Rahman ²²	8	PME ³²	1.0	Cut-off/Verlet	1.0	SPC ²⁷
Martini 2.2	0.02	v-rescale ²⁹	1	Parrinello-Rahman ²²	12	Reaction-Field ³³	1.1	Cut-off/Potential-shift	1.1	Martini ²⁸
Martini 3.0b3.2	0.02	v-rescale ²⁹	1	Parrinello-Rahman ²²	12	Reaction-Field ³³	1.1	Cut-off/Potential-shift	1.1	Martini ²⁸
Strah 2.1	0.02	Langevin	8	Berendsen ²¹	8	PME ³²	1.2	Cut-off	1.2	WT4 ³¹
Martini 2.2p	0.02	v-rescale ²⁹	1	Parrinello-Rahman ²²	12	Reaction-Field ³³	1.1	Cut-off/Potential-shift	1.1	POL ³⁰
Martini 2.3p	0.02	v-rescale ²⁹	1	Parrinello-Rahman ²²	12	Reaction-Field ³³	1.1	Cut-off/Potential-shift	1.1	POL ³⁰
Dry Martini (2014)	0.04	Langevin	4	Parrinello-Rahman ²²	8	Reaction-Field ³³	1.1	Cut-off/Potential-shift	1.1	—
Dry Martini (2016)	0.04	Langevin	4	Parrinello-Rahman ²²	8	Reaction-Field ³³	1.1	Cut-off/Potential-shift	1.1	—

Figure S2: Correlation plot for SOAP distance and Jensen-Shannon divergence for POPC. The correlation coefficient is $r = 0.58$, that confirms the compatibility between the two metrics.

Figure S3: Effective rank of the SOAP matrix in function of the cut off. We compared the rank considering all the 4 beads (orange line) or only the phosphate bead (blue line).

Figure S4: 2D projection of distances on the distance matrix via Multidimensional Scaling³⁴ of the AA and UA force fields.

References

- [1] Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comput. Phys. Commun.* **2014**, *185*, 604–613.
- [2] Bonomi, M. et al. *Nat. Methods* **2019**, *16*, 670–673.
- [3] Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. *J. Comput. Chem.* **2011**, *32*, 2319–2327.
- [4] Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. *SoftwareX* **2015**, *1*, 19–25.
- [5] Camley, B. A.; Lerner, M. G.; Pastor, R. W.; Brown, F. L. *J. Chem. Phys.* **2015**, *143*, 12B604_1.
- [6] Vögele, M.; Hummer, G. *J. Phys. Chem. B* **2016**, *120*, 8722–8732.
- [7] Vögele, M.; Köfinger, J.; Hummer, G. *Phys. Rev. Lett.* **2018**, *120*, 268104.
- [8] Bartók, A. P.; Kondor, R.; Csányi, G. *Phys. Rev. B* **2013**, *87*, 184115.
- [9] Behler, J. *J. Chem. Phys.* **2011**, *134*, 074106.
- [10] Rupp, M.; Tkatchenko, A.; Müller, K.-R.; von Lilienfeld, O. A. *Phys. Rev. Lett.* **2012**, *108*, 058301.
- [11] Faber, F.; Lindmaa, A.; von Lilienfeld, O. A.; Armiento, R. *Int. J. Quantum Chem.* **2015**, *115*, 1094–1101.
- [12] Huo, H.; Rupp, M. Unified Representation of Molecules and Crystals for Machine Learning. 2018.
- [13] De, S.; Bartók, A. P.; Csányi, G.; Ceriotti, M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 13754–13769.
- [14] Himanen, L.; Jäger, M. O. J.; Morooka, E. V.; Federici Canova, F.; Ranawat, Y. S.; Gao, D. Z.; Rinke, P.; Foster, A. S. *Comput. Phys. Commun.* **2020**, *247*, 106949.
- [15] Facco, E.; D’Errico, M.; Rodriguez, A.; Laio, A. *Sci. Rep.* **2017**, *7*, 1–8.
- [16] Erba, V.; Gherardi, M.; Rotondo, P. *Sci. Rep.* **2019**, *9*, 1–9.
- [17] Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y., et al. *J. Chem. Theory Comput.* **2016**, *12*, 405–413.
- [18] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- [19] Case, D. et al. AmberTools 2020. 2020.

- [20] Swails, J.; Hernandez, C.; Mobley, D.; Nguyen, H.; Wang, L.; Janowski, P. ParmEd. 2010.
- [21] Berendsen, H. J.; Postma, J. v.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- [22] Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- [23] Evans, D. J.; Holian, B. L. *J. Chem. Phys.* **1985**, *83*, 4069–4074.
- [24] Ferreira, T. M.; Coreta-Gomes, F.; Ollila, O. S.; Moreno, M. J.; Vaz, W. L.; Topgaard, D. *Phys. Chem. Chem. Phys.* **2013**, *15*, 1976–1989.
- [25] Samuli, O. O. H.; Tiago, F.; Daniel, T. MD simulation trajectory and related files for POPC bilayer (Berger model delivered by Tieleman, Gromacs 4.5). 2014; <https://doi.org/10.5281/zenodo.13279>.
- [26] Domański, J.; Stansfeld, P. J.; Sansom, M. S.; Beckstein, O. *J. Membr. Biol.* **2010**, *236*, 255–258.
- [27] Berendsen, H. J.; Postma, J. P.; van Gunsteren, W. F.; Hermans, J. *Inter-molecular forces*; Springer, 1981; pp 331–342.
- [28] Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
- [29] Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.
- [30] Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. *PLoS Comput. Biol.* **2010**, *6*, e1000810.
- [31] Darré, L.; Machado, M. R.; Dans, P. D.; Herrera, F. E.; Pantano, S. *J. Chem. Theory Comput.* **2010**, *6*, 3793–3807.
- [32] Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- [33] Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- [34] Kruskal, J. B. *Psychometrika* **1964**, *29*, 1–27.