ARTICLES
https://doi.org/10.1038/s41592-021-01098-3

nature methods

Check for updates

Martini 3: a general purpose force field for coarse-grained molecular dynamics

Paulo C. T. Souza[1,2] ✉, Riccardo Alessandri[1], Jonathan Barnoud[1,3], Sebastian Thallmair[1,4], Ignacio Faustino[1], Fabian Grünewald[1], Ilias Patmanidis[1], Haleh Abdizadeh[1], Bart M. H. Bruininks[1], Tsjerk A. Wassenaar[1], Peter C. Kroon[1], Josef Melcr[1], Vincent Nieto[2], Valentina Corradi[5], Hanif M. Khan[5,6], Jan Domański[7,8], Matti Javanainen[9,10], Hector Martinez-Seara[9], Nathalie Reuter[6], Robert B. Best[8], Ilpo Vattulainen[10,11], Luca Monticelli[2], Xavier Periole[12], D. Peter Tieleman[5], Alex H. de Vries[1] and Siewert J. Marrink[1] ✉

The coarse-grained Martini force field is widely used in biomolecular simulations. Here we present the refined model, Martini 3 (http://cgmartini.nl), with an improved interaction balance, new bead types and expanded ability to include specific interactions representing, for example, hydrogen bonding and electronic polarizability. The updated model allows more accurate predictions of molecular packing and interactions in general, which is exemplified with a vast and diverse set of applications, ranging from oil/water partitioning and miscibility data to complex molecular systems, involving protein–protein and protein–lipid interactions and material science applications as ionic liquids and aedamers.

The molecular dynamics simulation technique has become an indispensable tool in natural sciences, offering a spatio-temporal resolution unmatched by any experimental technique[1]. A major bottleneck of molecular dynamics is the limited time and length scales that are accessible. To overcome this limitation, coarse-grained (CG) models representing groups of atoms by effective beads, have achieved widespread use[2]. The Martini model is among the most popular CG models in the field of biomolecular simulation, due to its easy-to-use building block principle. Martini relies on a four-to-one mapping scheme (that is, on average four heavy atoms and associated hydrogens are mapped into one CG bead), and has been parametrized using a top-down approach with thermodynamic partitioning data as the main target[3,4]. Nonbonded interactions between neutral beads of Martini are solely described by Lennard-Jones potentials, while charged beads also include Coulombic interactions. The interaction strength of the Lennard-Jones potential (that is, its well depth) is used to discriminate between different levels of polarity of the CG beads. The model features four main classes of CG bead types, denoted C, N, P and Q representing nonpolar, intermediately polar, polar and charged chemical groups, respectively[4]. Sublabels are used to make a further distinction within a class in terms of degree of polarity or hydrogen donor/acceptor capabilities. In principle, all beads are the same size, denoted regular (R) beads. By way of exception, special small (S) beads were introduced to model ring-like compounds for which a four-to-one mapping scheme is inadequate[4]. To reproduce correct stacking and hydrogen-bonding distances between nucleotides, even smaller tiny (T) beads were found necessary[5]. Parametrization of the cross-interactions between S and T beads with R beads, however, was done on an ad hoc basis.

The Martini force field is used in a wide range of applications in diverse fields including structural biology[6–8], biophysics[9,10], biomedicine[11], nanotechnology[12,13] and materials design[14,15]. With its growing use, however, a number of shortcomings of the Martini model have recently been identified. One of the most important problems is the observation that certain molecules tend to interact too strongly. This has been reported for proteins and carbohydrates in solution, as well as for membrane embedded proteins[16–18]. The origin lies among others in small but systematic deviations in packing and intermolecular interactions[19]. Besides, the coverage of chemical space for broader applications was uneven, and in some cases, such as selectivity of nucleobase pairing[5,20], consistency was difficult to obtain given the limited bead types and sizes. To alleviate these problems, we undertook a rebalancing of all nonbonded interaction terms of the Martini model, including the addition of new beads and labels. The new version, called Martini 3, enables more accurate simulations of molecular systems in general. In this paper, we present the key features of Martini 3 combined with examples of new applications and improvements in relation to the previous Martini model.

Results

Reparametrization of the beads. In Martini 3, the parametrization strategy was based on the construction of prototype

[1]Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Material, University of Groningen, Groningen, the Netherlands. [2]Molecular Microbiology and Structural Biochemistry, UMR 5086 CNRS and University of Lyon, Lyon, France. [3]Intangible Realities Laboratory, University of Bristol, School of Chemistry, Bristol, UK. [4]Frankfurt Institute for Advanced Studies, Frankfurt am Main, Germany. [5]Centre for Molecular Simulation and Department of Biological Sciences, University of Calgary, Calgary, Alberta, Canada. [6]Department of Chemistry and Computational Biology Unit, University of Bergen, Bergen, Norway. [7]Department of Biochemistry, University of Oxford, Oxford, UK. [8]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, USA. [9]Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague, Czech Republic. [10]Computational Physics Laboratory, Tampere University, Tampere, Finland. [11]Department of Physics, University of Helsinki, Helsinki, Finland. [12]Department of Chemistry, Aarhus University, Aarhus C, Denmark. ✉e-mail: paulocts@gmail.com; s.j.marrink@rug.nl

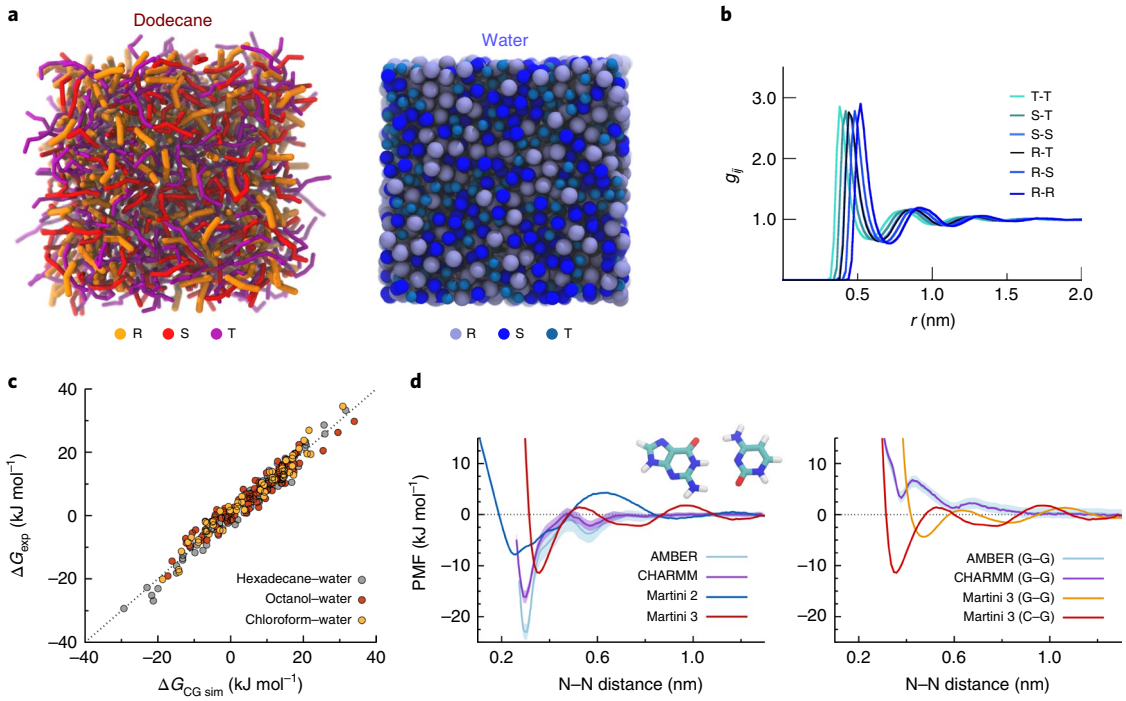

Fig. 1 | Rebalancing R, S and T beads. a, Snapshots of simulation boxes containing mixtures of dodecane and water in three resolutions. **b**, Radial distribution functions (g_{ij}) for all bead combinations in the multi-resolution mixture of water. **c**, Water–oil transfer free energies (ΔG) computed for around 260 data points using Martini 3. **d**, Hydrogen-bonding PMF between nucleobases. On the left, comparison between Martini 2 and 3 for the cytosine–guanine base pair. On the right, comparison of the cytosine–guanine (C–G) and guanine–guanine (G–G) base pairs using Martini 3. In both plots, CHARMM and AMBER atomistic data are also reported[5] for comparison. Errors are estimated with bootstrapping and displayed as transparent shades. In the case of Martini, errors are smaller than 0.1 kJ mol^{-1}, and hence are not visible on the graphs.

models of polar and nonpolar molecules in all three Martini resolutions. Self- (R-R, S-S, and T-T) and cross-interactions (R-S, R-T and S-T) of the different bead sizes were optimized to be well-balanced (Supplementary Notes). In terms of chemical types, the beads were separated into three blocks: organic, ions and water (Supplementary Notes). The organic (containing P, N and C beads) and ion (Q beads) blocks have been subjected to independent parametrizations, where different trends in self-interaction, solvation and transfer free energy on bead size change were included (Supplementary Notes). In contrast to the previous version, water is defined as a separate bead type (called W), which enables optimization of water properties independently from other targets; for example, the freezing of water at room temperature (a problem sometimes encountered with the previous water model) no longer occurs. In addition, it is available in three different sizes as well (Supplementary Notes). Together with this optimization strategy, the new Martini 3 model also features a fully revised interaction matrix (Supplementary Notes) and new intermediate interaction levels, added to smoothen the transition between chemical types (Supplementary Notes). Bead assignment and validation of the models were not only based on experimental transfer free energies, but also included solvent miscibility data (Supplementary Notes and Supplementary Results) and a series of benchmark tests, ranging from structural properties of bilayers to dimerization potentials of mean force (PMF) of proteins (Methods, Supplementary Notes and Supplementary Results).

The improved interaction balance between regular and smaller bead types is illustrated by the close to ideal mixing behavior of pure solvents composed of molecules mapped at different resolutions (Fig. 1a). Integration of radial distribution functions, defined as Kirkwood–Buff integrals (G_{ij}), are used here to quantify the degree of miscibility of the multi-resolution liquid water model (Fig. 1b). Theoretically, pair differences in Kirkwood–Buff integrals (ΔG_{ij}) should be equal to zero for all i,j pairs in ideal mixtures[21,22] while real mixtures that closely approach ideal behavior (such as benzene–toluene) show values around ± 1 cm^3 mol^{-1} (ref. [22]). Our multi-resolution water model shows $\Delta G_{ij} \approx -0.3$ cm^3 mol^{-1}, indicating that the balance achieved with the new parametrization closely captures an ideal mixing behavior.

The accuracy of CG models containing ring or branched fragments, which rely heavily on smaller bead types, is also greatly increased in Martini 3. For example, the binary mixing behavior of various solvents (Supplementary Results) and the transfer free energies of many linear, branched and ring-like compounds (Fig. 1c and Supplementary Results) are now in very good agreement with experimental data. The mean absolute error of transfer free energies compared to the experimental data is 2 kJ mol^{-1}, with 86% of the molecules presenting errors lower than 3 kJ mol^{-1}.

Another benefit of the recalibrated interactions is the disappearance of the artificially large desolvation free energy barriers that contribute to slow dissociation processes of the previous Martini 2 models. The problem that was initially observed in dimerization of nucleobases[5,19] is thus solved, as highlighted by the comparison of Martini 2 and Martini 3 PMFs between cytosine and guanine (left panel of Fig. 1d). Note that there is room for further improvement, as the free energy minima of the CG PMF profiles with Martini 3 are shifted relative to the all-atom profiles because the bead sizes representing nitrogen-containing groups are not optimal to reproduce hydrogen-bonding distances. In addition, the difference between C–G and G–G base pairs is not as large as in the atomistic case (roughly 20 kJ mol^{-1}). However, it is still large enough (roughly 8 kJ mol^{-1}) to provide specificity.

The proper balancing of R, S and T beads in Martini 3 also indicates that the mapping choice of an arbitrary molecule to its Martini representation is now better defined. S and T beads are not only suited to represent ring-like compounds, but also used for cases involving 3-to-1 and 2-to-1 mapping of linear and branched chemical groups (Supplementary Notes).

Covering the chemical space with new beads and labels. Together with a thorough revision of the interaction strengths, in Martini 3, we extend the number of chemical bead types and the ability to modify the bead properties depending on the chemical details of the underlying moieties. Each P, N and C class now has six bead types with different degrees of polarity, which enables a more precise definition of different chemical groups by assigning them to certain bead types. Additionally, we introduce a new X-class of beads to model halo-compounds (Supplementary Notes). In the previous version of Martini, some of the bead types were already subclassified according to their ability to act as hydrogen bond donor, acceptor or both. This property can be now attributed to all bead types of an intermediate or polar nature (N or P class). The effective interaction strength between donor and acceptor pairs is increased, whereas donor–donor and acceptor–acceptor pairs are weakened (Supplementary Notes). For example, Martini 3 correctly reproduces the trends in hydrogen bond-based pairing of nucleobases[5,20] without the use of special-purpose beads specifically for nucleobases (right panel of Fig. 1d). Note that chemical groups that can act as both donor and acceptor at the same time are always represented by the pure beads of the P and N class in Martini 3.

Next to the fine-tuning based on hydrogen-bonding capabilities, we introduce the possibility to change the interactions based on the electronic polarizability. Depending on inductive or conjugate effects caused by chemical functionalization, nonpolar molecules can be polarized; that is, they can acquire an electron-donor (or 'enriched', label 'e') or electron-acceptor (or 'vacancy', label 'v') character, which can promote preferential interactions. Polarizable groups in Martini 3 can be distinguished through the label 'e/v' that can only be applied to the C- and X-class. A nice example of their application is the strong and specific interaction between electron-donor and electron-acceptor aromatic rings in aedamers, a class of molecules that have been studied extensively in the context of biomimetic folding and self-assembly[23,24]. The use of 'e/v' allows Martini 3 to capture the preferential interaction between 1,5-dialkoxynaphthalene (DAN) and naphthalene diimide (NDI) (left panel of Fig. 2a) experimentally observed via nuclear magnetic resonance (NMR) titration[23] and atomistic simulation data. Self-assembly of amide-linked tetramers shows preferential formation of alternating stacks of DAN and NDI, which is also measured by NMR and isothermal titration calorimetry investigations[24]. On top of hydrogen bonding and electron polarization labels, all beads can have their self-interaction fine-tuned by other sublabels (as further described in the Supplementary Notes).

Chemical groups carrying monovalent charges +1/−1 are represented in Martini by the class of Q beads (Supplementary Notes). The original Martini model only considers monovalent ions, and was solely optimized for regular bead sizes that represented small ions and their first hydration shell. In Martini 3, charged groups can have R, S or T sizes. The tiny size category allows modeling of small, bare ions, enabling applications that involve ion binding where (part of) the hydration shell is lost. This feature is exemplified by the binding of sodium ions (represented by a charged tiny bead) to a buried small cavity localized in the core of the adenosine A_{2A} receptor (Fig. 2b). X-ray crystallographic[25] and ligand binding assays[26] confirm the importance of sodium ions for the structure and for the allosteric modulation of the A_{2A} receptor. Note that an extensive validation of the lipid models in Martini 3 was performed to allow simulations of transmembrane and peripheral membrane proteins (Supplementary Results).

In addition to the smaller sizes, the Q class was also expanded to five bead types, following the classical Hofmeister series trend[27,28] (Supplementary Notes and Supplementary Results). At one extreme, the Q5 bead may be used to represent hard monovalent ions with the smallest polarizability, for example inorganic ions such as $R_2PO_4^-$. At the other end of the Martini–Hofmeister series, the Q1 type models polarizable soft monovalent ions, such as $N(CH_3)_4^+$, and includes ion–π interactions via the Lennard-Jones potential. Such differences in behavior of the different Q-bead types are exemplified by molecular dynamics simulations of the anion transfer between aqueous solutions and organophosphonium-based ionic liquids (Fig. 2c and Supplementary Results). Harder ions such as Cl^- (modeled as TQ5 with −1 charge) tend to stay in the water phase, together with Na^+ ions (TQ5$^+$ bead). In contrast, softer ions such as ClO_4^- (Q2$^-$ bead) can exchange with Br^- (SQ4$^-$ bead) or (partially) PF_6^- (Q1$^-$ bead) from the ionic liquid phase. In the case of the biphasic system using trihexyltetradecylphosphonium bromide ($[P_{66614}][Br]$), direct comparison to experimental data shows good agreement for the anion transfer trends[28,29]. The new Q-bead types also impact biologically relevant systems, as exemplified by preferential cation–π interaction between choline groups (Q1$^+$ bead) of phosphatidylcholine lipids and aromatic residues of the *Bacillus thuringiensis* phosphatidylinositol-specific phospholipase C (*Bt*PI-PLC). In the previous version, such specific interactions between soft ions and aromatic molecules were solely included in the recently updated polarizable Martini implementation[30]. However, in Martini 3, the different Q-bead types allow easier (implicit) incorporation of such interactions without the need for additional partial charges.

On top of the new chemical types, all Q beads can use the hydrogen-bonding labels (called in this case 'p/n'). They represent organic charged molecules or fragments, such as R-CH_2-COO^- and R-CH_2-NH_3^+, and also introduce modifications in the Hofmeister trends of the pure Q beads (Supplementary Notes). Positively charged hydrogen donors ('p' label) interact more strongly with nonpolar beads, as expected in cation–π interactions. On the other hand, negatively charged hydrogen acceptors ('n' label), have stronger interactions with neutral polar beads, mimicking the stronger hydrogen bonds with anions. To complete the ion block, we explicitly include a new D bead for divalent ions (such as Ca^{2+}), which are typically hard ions.

Improving packing and protein–protein interactions. Another change in philosophy with respect to the previous Martini models is the definition of bonded interactions. Instead of using the center of mass of the mapped chemical groups to define the geometry of the molecule, we now use a size-shape concept aimed at preserving the volume of molecules in comparison to all-atom reference structures. This choice and the proper use of Martini 3 bead sizes (Supplementary Notes) lead to more realistic molecular packing. As a consequence, the hydration of protein pockets and channels is improved, as illustrated by the Fragaceatoxin C (FraC) nanopore inserted in a lipid bilayer (Fig. 3a). The pore of FraC remains open over the whole trajectory in Martini 3, as indicated by X-ray crystallography[31] and electro-osmotic flow assays[32], while in Martini 2 it is closed.

Another example of accurate packing is the stacking predictions of thiophene derivatives in bulk heterojunction solar cells composed of poly(3-hexyl-thiophene) (P3HT) and phenyl-C61-butyric acid methyl ester (PCBM) (Fig. 3b). The morphology of these organic solar cells is a determinant for high-efficiency devices[33]. The scattering profiles computed with Martini 3 show improved agreement with Martini 3 in relation to P3HT lamellar (peak around $q \cong 0.45$ Å^{-1}) and stacking ($q \cong 1.65$ Å^{-1}) experimental distances[33,34].

The use of bonds based on molecular volume and the appropriate choice of chemical bead types, sizes and labels also controls the interaction density of the model, which has an important impact

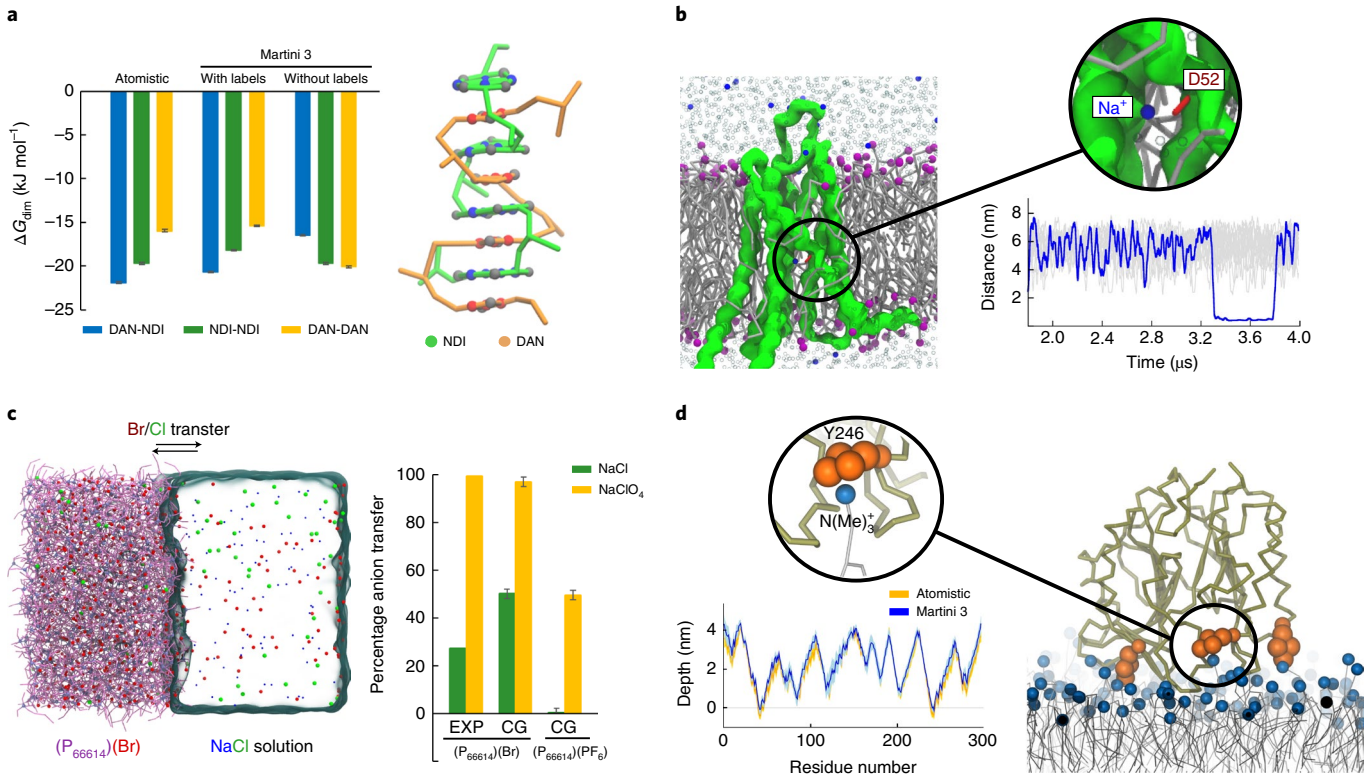

Fig. 2 | New chemical bead types, sublabels and applications. a, Self-assembly of aedamers. The left panel shows the dimerization free energies (ΔG_{dim}) of pegylated monomers of DAN and NDI. Errors are estimated with bootstrapping. The right panel shows the self-assembled duplex dimer formed by amide-linked tetramers of NDI (green) and DAN (orange). **b**, As indicated by X-ray crystallography[25], sodium ions (charged TQ5 bead) can bind to a buried small cavity in the core of the adenosine A_{2A} receptor. **c**, Charged Q beads in Martini 3 follow the classical Hofmeister series, as exemplified by the anion transfer between salt aqueous solutions and organophosphonium-based ionic liquids (right panel). Errors in the average anion transfer percentage are estimated by block averaging. **d**, Preferential cation–π interaction between choline groups (Q1 bead) of phosphatidylcholine lipids and aromatic residues of the *B. thuringiensis* phosphatidylinositol-specific phospholipase C (*Bt*PI-PLC). The depth of insertion of each amino acid of *Bt*PI-PLC is in very good agreement with the insertion obtained from an atomistic molecular dynamics simulation[30].

on the strength of collective interactions between molecules[19]. To test to what extent the changes in nonbonded and bonded interactions reduce the over-estimated aggregation of proteins, we performed extensive simulations comprising solutions of soluble proteins as well as membrane embedded proteins. These systems were simulated under conditions in which proteins do not aggregate and, preferentially, stay as monomers. For soluble proteins (Fig. 3c), qualitative tests were performed with the headpiece domain of chicken villin[35], and the modified and mutated cellulose-binding domain from *Cellulomonas fimi* (EXG-CBM), which is an example of a protein completely free of charged side chains that can maintain solubility, stability and function[36]. Trends are well captured in Martini 3, with both proteins mainly staying as monomers in pure water (with only counter-ions to neutralize the system in the case of villin). The villin headpiece showed salting-in behavior (that is, less aggregation) under addition of 0.4 M of NaCl, which was also observed for certain soluble proteins at low ionic strengths[37]. On the other hand, EXG-CBM only showed salting-out behavior (that is, more aggregation), which was expected based on experimental data[36]. In contrast, both proteins aggregate in Martini 2, forming a single and stable aggregate during the simulation.

Polyleucine (K_2-L_{26}-K_2) was selected to evaluate the aggregation propensity in membranes. Experimental evidence with this transmembrane protein model indicates a preference for the monomeric state in a bilayer environment[38–40]. Both Martini 2 and 3 show that the hydrophobic mismatch between transmembrane length and membrane thickness can play a role in the aggregation, with Martini 3 showing a higher percentage of the monomeric state (Fig. 3d). To quantitatively evaluate the strength of protein–protein interactions in a membrane environment, we also considered the dimerization of four selected transmembrane helices: the transmembrane domains of the receptor tyrosine kinases EphA1 and ErbB1; the red blood cell protein glycophorin A (GpA); as well as the well-known model peptide WALP23 (left panel of Fig. 3e). For EphA1 and ErbB1, experimental dimerization free energies in a membrane environment have been estimated using Förster resonance energy transfer (FRET)[41,42]. For GpA, dimerization free energies range from around −15 kJ mol^{-1} (in various cell membrane environments)[43,44] to −31.5 kJ mol^{-1} (GALLEX assay in *E. coli* inner membranes)[45,46] or −50.6 kJ mol^{-1} (steric trap in 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayers)[47]. WALP peptides have been characterized thoroughly during the past two decades, including their self-association[48]. For each one of the four peptide dimers, we compared experimental dimerization free energies with the free energies predicted by the Martini 2 and Martini 3 models. Martini 3 shows not only to be able to capture the correct trends, but also to quantitatively reproduce the experimental affinities. The binding mode also becomes improved as highlighted for GpA (right panel of Fig. 3e). The GpA homodimer structure with Martini 3 closely resembles experimental results obtained with NMR spectroscopy and crystallography[49–51].

In summary, for both soluble and transmembrane proteins, we observed that the Martini 3 models are in much better agreement with experimental data than before. Another advantage of the current

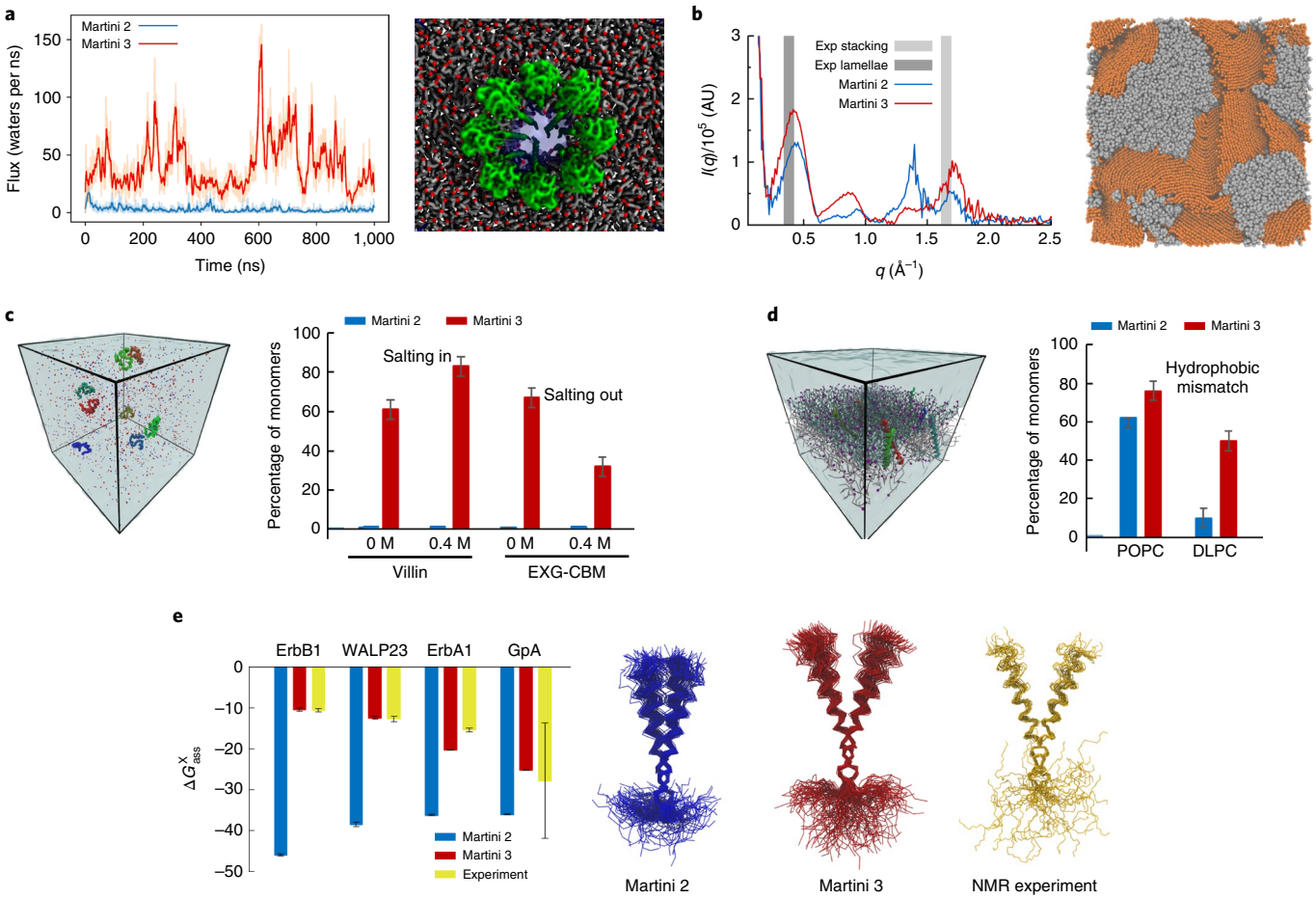

Fig. 3 | Improving packing, cavities and reducing protein stickiness. a, Hydration of FraC nanopore inserted in a lipid bilayer. **b**, Scattering profiles and a Martini 3 snapshot of a bulk heterojunction morphology of P3HT (in orange) and PCBM (in gray) formed after solvent evaporation and annealing simulations. *I(q)* corresponds to scattering intensity and *q* is the reciprocal space vector. **c**, Aggregation levels of the soluble proteins villin headpiece and the modified EXG-CBM in different salt concentrations. **d**, Aggregation levels of polyleucine helices in POPC and 1,2-dilauroyl-sn-glycero-3-phosphocholine (DLPC) bilayers. Errors in the average monomer percentage of **c** and **d** are estimated by block averaging. **e**, Dimerization of transmembrane helices. The left panel shows a comparison between experimental and calculated values for the mole fraction standard Gibbs free energy of dimerization (ΔG^{X}_{ass}) of the following transmembrane protein domains: ErbB1, EphA1, WALP23 and GpA. Simulation errors are estimated with bootstrapping while experimental data were obtained from the literature[41–48]. In the case of GpA, error was estimated by the mean absolute error of four independent experimental data[43–47]. A comparison between experimental and simulated binding modes of GpA is highlighted in the right panel. The experimental structure was taken from solution NMR in micelles (Protein Data Bank accession code 1AFO)[49]. Near identical experimental structures were obtained by solid-state NMR in nanodiscs and X-ray crystallography in a lipid cubic phase[49–51].

Martini 3 protein model is that the default bead type representing the protein backbone in Martini 3 (a regular P2 bead) no longer depends on the secondary structure. In addition, the representation of protein flexibility can now be improved by the use of Gō-like models[52].

Discussion

In this paper, we have described the new version of the Martini force field, which shows numerous improvements in relation to the previous version. However, inherent limitations to the process of coarse graining, related to transferability and representability problems[53–55] are still part of the model. An important drawback is the limited structural detail that is a consequence of representing multiple atoms with isotropic interaction sites. This is most noticeable for the Martini water model, which represents four water molecules with a single Lennard-Jones site and will certainly not capture any of the higher order structural correlations of real water. The role of explicit water in a CG model such as Martini is mostly to provide a good solvent for polar compounds resulting in realistic partitioning. For applications requiring finer details, structure-based CG models are more suitable[56,57]. Another fundamental limitation is the entropy–enthalpy compensation. The loss of internal degrees of freedom for groups of atoms represented by a CG bead inevitably reduces the entropy of the system. Since the Martini force field is based on reproducing free energies, this requires a concomitant reduction in the enthalpy. As consequence, inaccurate entropy–enthalpy balance affects the temperature dependence of several properties and reduces the transferability to different state points. To probe transferability, we performed temperature-dependent calculations for a number of solvent systems as well as lipid membranes (Supplementary Results). Temperature-dependent properties, such as the heat expansion coefficient and heat capacity of water and n-octane, are very well captured by Martini 3, but this is not true for the hydrophobic effect that shows the opposite trend with respect to atomistic models, in line with previous findings[58]. Note that the use of higher-resolution S or T particles does not remedy this problem, as these bead types were parameterized mainly to be compatible with the regular (R type) beads and should be used primarily to represent parts of the system that cannot be adequately

mapped with R particles. Potential improvements with respect to the temperature transferability of our model could make use of environment dependent potentials[59] or CG beads with embedded sites, such as the polarizable water models[60,61] where incorporation of quadrupole moment might be required[58]. Bottom-up CG models that are derived with minimization of the information loss[54] as parameterization target might also perform better. For a more in depth discussion of these and related issues with respect to the Martini coarse-graining philosophy, we refer to previous papers[62,63].

Keeping these limitations in mind, Martini 3 offers a versatile and easy-to-use generic force field to simulate molecular processes at a semi-quantitative level of accuracy. In relation to the previous model, the excessive over-estimated aggregation[19] is substantially reduced. We expect that Martini 3 will allow for more realistic predictions of protein interactions, as well as more accurate simulations of molecular systems in general. The increased number of bead types and interaction levels makes the model even more versatile, covering a larger part of chemical space with appropriate building blocks. Based on this foundation, further optimizations are currently ongoing, including the use of Gō-potentials to alleviate limitations of protein conformational flexibility, a reoptimization of the bonded potentials of lipids and other biomolecular classes, as well as a compatible polarizable water model for applications requiring more realistic screening of electrostatic interactions. Finally, we foresee new application horizons for the Martini model in the field of materials science[64,65] and high-throughput drug design[66].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests and statements of data and code availability are available at https://doi.org/10.1038/s41592-021-01098-3.

Received: 16 June 2020; Accepted: 22 February 2021;
Published online: 29 March 2021

References

1. Bottaro, S. & Lindorff-Larsen, K. Biophysical experiments and biomolecular simulations: a perfect match? *Science* **361**, 355–360 (2018).
2. Ingólfsson, H. I. et al. The power of coarse graining in biomolecular simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4**, 225–248 (2014).
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. 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).
5. Uusitalo, J. J., Ingólfsson, H. I., Akhshi, P., Tieleman, D. P. & Marrink, S. J. Martini coarse-grained force field: extension to DNA. *J. Chem. Theory Comput.* **11**, 3932–3945 (2015).
6. Abellón-Ruiz, J. et al. Structural basis for maintenance of bacterial outer membrane lipid asymmetry. *Nat. Microbiol.* **2**, 1616–1623 (2017).
7. Yen, H. Y. et al. PtdIns(4,5)P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **559**, 423–427 (2018).
8. 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).
9. Vögele, M., Köfinger, J. & Hummer, G. Hydrodynamics of diffusion in lipid membrane simulations. *Phys. Rev. Lett.* **120**, 268104 (2018).
10. Agostino, M. D., 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).
11. Jeena, M. T. et al. Mitochondria localization induced self-assembly of peptide amphiphiles for cellular dysfunction. *Nat. Commun.* **8**, 26 (2017).
12. Jiang, Z. et al. Subnanometre ligand-shell asymmetry leads to Janus-like nanoparticle membranes. *Nat. Mater.* **14**, 912–917 (2015).
13. Maingi, V. et al. Stability and dynamics of membrane-spanning DNA nanopores. *Nat. Commun.* **8**, 14784 (2017).
14. Frederix, P. W. J. M. et al. Exploring the sequence space for (tri-)peptide self-assembly to design and discover new hydrogels. *Nat. Chem.* **7**, 30–37 (2015).
15. Bochicchio, D., Salvalaglio, M. & Pavan, G. M. Into the dynamics of a supramolecular polymer at submolecular resolution. *Nat. Commun.* **8**, 147 (2017).
16. Stark, A. C., Andrews, C. T. & Elcock, A. H. Toward optimized potential functions for protein-protein interactions in aqueous solutions: osmotic second virial coefficient calculations using the MARTINI coarse-grained force field. *J. Chem. Theory Comput.* **9**, 4176–4185 (2013).
17. Javanainen, M., Martinez-Seara, H. & Vattulainen, I. Excessive aggregation of membrane proteins in the Martini model. *PLoS ONE* **12**, e0187936 (2017).
18. Schmalhorst, P. S., Deluweit, F., Scherrers, R., Heisenberg, C.-P. & Sikora, M. Overcoming the limitations of the MARTINI force field in simulations of polysaccharides. *J. Chem. Theory Comput.* **13**, 5039–5053 (2017).
19. Alessandri, R. et al. Pitfalls of the Martini model. *J. Chem. Theory Comput.* **15**, 5448–5460 (2019).
20. Uusitalo, J. J., Ingólfsson, H. I., Marrink, S. J. & Faustino, I. Martini coarse-grained force field: extension to RNA. *Biophys. J.* **113**, 246–256 (2017).
21. Ben-Naim, A. *Molecular Theory of Solutions* (Oxford Univ. Press, 2006).
22. Ploetz, E. A., Bentenitis, N. & Smith, P. E. Kirkwood–Buff integrals for ideal solutions. *J. Chem. Phys.* **132**, 164501 (2010).
23. Zych, A. J. & Iverson, B. L. Synthesis and conformational characterization of tethered, self-complexing 1,5-dialkoxynaphthalene/1,4,5,8-n aphthalenetetracarboxylic diimide systems. *J. Am. Chem. Soc.* **122**, 8898–8909 (2000).
24. Gabriel, G. J. & Iverson, B. L. Aromatic oligomers that form hetero duplexes in aqueous solution. *J. Am. Chem. Soc.* **124**, 15174–15175 (2002).
25. Liu, W. et al. Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **337**, 232–236 (2012).
26. Gao, Z. G. & Ijzerman, A. P. Allosteric modulation of A(2A) adenosine receptors by amiloride analogues and sodium ions. *Biochem. Pharmacol.* **60**, 669–676 (2000).
27. Okur, H. I. et al. Beyond the Hofmeister series: ion-specific effects on proteins and their biological functions. *J. Phys. Chem. B* **121**, 1997–2014 (2017).
28. Dupont, D., Depuydt, D. & Binnemans, K. Overview of the effect of salts on biphasic ionic liquid/water solvent extraction systems: anion exchange, mutual solubility, and thermomorphic properties. *J. Phys. Chem. B* **119**, 6747–6757 (2015).
29. Naert, P., Rabaey, K. & Stevens, C. V. Ionic liquid ion exchange: exclusion from strong interactions condemns cations to the most weakly interacting anions and dictates reaction equilibrium. *Green Chem.* **20**, 4277–4286 (2018).
30. Khan, H. M. et al. Capturing choline-aromatics cation–π interactions in the MARTINI force field. *J. Chem. Theory Comput.* **16**, 2550–2560 (2020).
31. Tanaka, K., Caaveiro, J. M. M., Morante, K., González-Manãs, J. M. & Tsumoto, K. Structural basis for self-assembly of a cytolytic pore lined by protein and lipid. *Nat. Commun.* **6**, 6337 (2015).
32. Huang, G., Willems, K., Soskine, M., Wloka, C. & Maglia, G. Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with FraC nanopores. *Nat. Commun.* **8**, 935 (2017).
33. Alessandri, R., Uusitalo, J. J., De Vries, A. H., Havenith, R. W. A. & Marrink, S. J. Bulk heterojunction morphologies with atomistic resolution from coarse-grain solvent evaporation simulations. *J. Am. Chem. Soc.* **139**, 3697–3705 (2017).
34. Chiu, M. Y., Jeng, U. S., Su, C. H., Liang, K. S. & Wei, K. H. Simultaneous use of small- and wide-angle X-ray techniques to analyze nanometerscale phase separation in polymer heterojunction solar cells. *Adv. Mater.* **20**, 2573–2578 (2008).
35. Petrov, D. & Zagrovic, B. Are current atomistic force fields accurate enough to study proteins in crowded environments? *PLoS Comput. Biol.* **10**, e1003638 (2014).
36. Højgaard, C. et al. A soluble, folded protein without charged amino acid residues. *Biochemistry* **55**, 3949–3956 (2016).
37. Ruckenstein, E. & Shulgin, I. L. Effect of salts and organic additives on the solubility of proteins in aqueous solutions. *Adv. Colloid Interface Sci.* **123–126**, 97–103 (2006).
38. Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T. & Engelman, D. M. Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat. Struct. Biol.* **7**, 154–160 (2000).
39. Zhou, F. X., Merianos, H. J., Brunger, A. T. & Engelman, D. M. Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl Acad. Sci. USA* **98**, 2250–2255 (2001).
40. Grau, B. et al. The role of hydrophobic matching on transmembrane helix packing in cells. *Cell Stress* **1**, 90–106 (2017).
41. Chen, L., Merzlyakov, M., Cohen, T., Shai, Y. & Hristova, K. Energetics of ErbB1 transmembrane domain dimerization in lipid bilayers. *Biophys. J.* **96**, 4622–4630 (2009).
42. Artemenko, E. O., Egorova, N. S., Arseniev, A. S. & Feofanov, A. V. Transmembrane domain of EphA1 receptor forms dimers in membrane-like environment. *Biochim. Biophys. Acta* **1778**, 2361–2367 (2008).
43. Sarabipour, S. & Hristova, K. Glycophorin A transmembrane domain dimerization in plasma membrane vesicles derived from CHO, HEK 293T, and A431 cells. *Biochim. Biophys. Acta - Biomembr.* **1828**, 1829–1833 (2013).

44. Chen, L., Novicky, L., Merzlyakov, M., Hristov, T. & Hristova, K. Measuring the energetics of membrane protein dimerization in mammalian membranes. *J. Am. Chem. Soc.* **132**, 3628–3635 (2010).
45. Nash, A., Notman, R. & Dixon, A. M. De novo design of transmembrane helix–helix interactions and measurement of stability in a biological membrane. *Biochim. Biophys. Acta - Biomembr.* **1848**, 1248–1257 (2015).
46. Finger, C. et al. The stability of transmembrane helix interactions measured in a biological membrane. *J. Mol. Biol.* **358**, 1221–1228 (2006).
47. Hong, H., Blois, T. M., Cao, Z. & Bowie, J. U. Method to measure strong protein–protein interactions in lipid bilayers using a steric trap. *Proc. Natl Acad. Sci. USA* **107**, 19802–19807 (2010).
48. Sparr, E. et al. Self-association of transmembrane α-helices in model membranes: importance of helix orientation and role of hydrophobic mismatch. *J. Biol. Chem.* **280**, 39324–39331 (2005).
49. MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. Transmembrane helix dimer: structure and implications. *Science* **276**, 131–133 (1997).
50. Trenker, R., Call, M. E. & Call, M. J. Crystal structure of the glycophorin A transmembrane dimer in lipidic cubic phase. *J. Am. Chem. Soc.* **137**, 15676–15679 (2015).
51. Domański, J., Sansom, M. S. P., Stansfeld, P. J. & Best, R. B. Balancing force field protein–lipid interactions to capture transmembrane helix–helix association. *J. Chem. Theory Comput.* **14**, 1706–1715 (2018).
52. Souza, P. C. T., Thallmair, S., Marrink, S. J. & Mera-Adasme, R. An allosteric pathway in copper, zinc superoxide dismutase unravels the molecular mechanism of the G93A amyotrophic lateral sclerosis-linked mutation. *J. Phys. Chem. Lett.* **10**, 7740–7744 (2019).
53. Brini, E. et al. Systematic coarse-graining methods for soft matter simulations-a review. *Soft Matter* **9**, 2108–2119 (2013).
54. Foley, T. T., Shell, M. S. & Noid, W. G. The impact of resolution upon entropy and information in coarse-grained models. *J. Chem. Phys.* **143**, 243104 (2015).
55. Wagner, J. W., Dama, J. F., Durumeric, A. E. P. & Voth, G. A. On the representability problem and the physical meaning of coarse-grained models. *J. Chem. Phys.* **145**, 044108 (2016).
56. Wörner, S. J., Bereau, T., Kremer, K. & Rudzinski, J. F. Direct route to reproducing pair distribution functions with coarse-grained models via transformed atomistic cross correlations. *J. Chem. Phys.* **151**, 244110 (2019).
57. Noid, W. G., Chu, J. W., Ayton, G. S. & Voth, G. A. Multiscale coarse-graining and structural correlations: connections to liquid-state theory. *J. Phys. Chem. B.* **111**, 4116–4127 (2007).
58. Wu, Z., Cui, Q. & Yethiraj, A. Driving force for the association of hydrophobic peptides: the importance of electrostatic interactions in coarse-grained water models. *J. Phys. Chem. Lett.* **2**, 1794–1798 (2011).
59. Jin, J., Yu, A. & Voth, G. A. Temperature and phase transferable bottom-up coarse-grained models. *J. Chem. Theory Comput.* **16**, 6823–6842 (2020).
60. Yesylevskyy, S. O., Schäfer, L. V., Sengupta, D. & Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **6**, e1000810 (2010).
61. Michalowsky, J., Schäfer, L. V., Holm, C. & Smiatek, J. A refined polarizable water model for the coarse-grained MARTINI force field with long-range electrostatic interactions. *J. Chem. Phys.* **146**, 054501 (2017).
62. Marrink, S. J. & Tieleman, D. P. Perspective on the Martini model. *Chem. Soc. Rev.* **42**, 6801–22 (2013).
63. Bruininks, B. M. H., Souza, P. C. T. & Marrink, S. J. in *Biomolecular Simulations: Methods and Protocols* (eds Bonomi, M. & Camilloni, C.) 105–127 (Humana Press Inc., 2019).
64. Liu, J. et al. Enhancing molecular n-type doping of donor-acceptor copolymers by tailoring side chains. *Adv. Mater.* **30**, 1704630 (2018).
65. Vazquez-Salazar, L. I., Selle, M., de Vries, A., Marrink, S. J. & Souza, P. C. T. Martini coarse-grained models of imidazolium-based ionic liquids: from nanostructural organization to liquid–liquid extraction. *Green Chem.* **22**, 7376–7386 (2020).
66. Souza, P. C. T. et al. Protein–ligand binding with the coarse-grained Martini model. *Nat. Commun.* **11**, 3714 (2020).

Methods

CG models. CG mappings of small molecules were initially inspired by the standard Martini 2 models, when they were available. Due to the well-balanced properties of the regular (R), small (S) and tiny (T) beads in Martini 3, the CG models now follow more specific rules for mapping. For instance, over-representing 3-to-1 or 2-to-1 fragments by the usage of R beads is always avoided. Aromatic rings without substituents are composed of T beads and, in case of substituents, S beads are used. Aliphatic rings without substituents are usually based on S beads, which better reproduce their molecular shape. More technical details about the mapping rules and bead types used are given in the Supplementary Notes. As in the previous version of Martini[5,20,67–69], bonded parameters are based on atomistic simulations or high-resolution experimental data. The main difference in Martini 3 lies in the protocol to derive bond lengths, which are now based on matching overall volume and shape of the molecules (Supplementary Notes). In this spirit, the bonded parameters of the protein models were also slightly modified from the standard Martini 2.2 values[68,70], including the addition of side chain corrections[71], based on experimental reference structures. Backbone bead types do not depend on the secondary structure anymore, but are now represented by P2 beads, except for proline (SP1a), alanine (SP2, with an additional bead for the side chain) and glycine (SP1). Adapted versions of Gō-like models[72] or Elastic Networks[73] were used to maintain the tertiary protein structure. All CG protein models were constructed using Martinize2, described in Supplementary Codes. Lipid mapping was inspired by the previous Martini model[4,74–76], but now following the Martini 3 rules for mapping and also with adaptations in the bonded parameters inspired by the 'extensible model' of Carpenter et al.[77].

General setup for CG molecular dynamics simulations and analysis. Settings for the CG simulations followed, in general, the 'new' Martini set of simulation parameters[78] using the leap-frog algorithm[79] for integrating the equations of motion. The Verlet neighbor search algorithm[80] is used to update the neighbor list every 20 steps with a buffer tolerance of 0.005 kJ mol^{-1} ps^{-1}. For the Lennard-Jones terms, we used a cutoff scheme with a value of 1.1 nm and the Verlet cutoff scheme[81] for the potential-shift. Long range electrostatic interactions were treated with reaction field[82] or particle mesh Ewald[83], with relative permittivity set to $\varepsilon_r = 15$ and a cutoff value of 1.1 nm. Reaction field was used for most of the systems, except the ones explicitly mentioning particle mesh Ewald. Periodic boundary conditions were used in all three dimensions. For the production simulations, the velocity rescaling thermostat[84] (coupling time constant of 1.0 ps) and the Parrinello–Rahman barostat[85] (coupling time constant of 12.0 ps) were employed to maintain temperature and pressure, respectively. Except for equilibration runs, a time step of 20 fs was used for all systems. CG simulation settings are available as input files for GROMACS on the Martini portal http://cgmartini.nl. GROMACS 2016.x and 2018.x were used to run all the molecular dynamics simulations[86,87]. For automated running and managing the Martini 3 simulations, we provide an updated version of Martinate[88,89], described in Supplementary Codes. All the analysis were performed using gmx analysis tools (GROMACS 2016 and 2018)[86,87], VMD v.1.9.4a12 (ref. [90]), xmgrace (v.5.1.25) and MDAnalysis[91]. The graphs were plotted using Excel 2016, xmgrace (v.5.1.25) and gnuplot (v.5.2). Figures were compiled using VMD v.1.9.4a12 and Inkscape v.1.1.

Parameter calibration, tests and validation. To parametrize the Lennard-Jones parameters of single beads and also test the Martini 3 CG models, many molecular systems and methods were used in this work. The overall iterative approach was not based in rigorous separation of calibration and validation groups. As Martini is based on pair interactions, it is hard to find simple systems that cover enough points in the interaction matrix for all bead size combinations. So, complex systems are not only used for validation but can also be part of the calibration. The tests performed were separated in 'tiers', which represent systems with different level of complexity. In tier 0, isolate beads and simple molecules are mainly used for calibration of Lennard-Jones parameters, with balance of different bead sizes and thermodynamics data (for example, liquid–liquid partitioning and miscibility) used as the main targets. In the intermediate tier 1, bilayer properties are checked, together with qualitative tests, and applied to systems such as soluble and transmembrane proteins. These qualitative tests are designed as 'yes-or-no' questions to evaluate the overall quality of the force field. At the same time, some points in the interaction matrix were also tested and fine-tuned here. In the final tier 2, quantitative tests involving complex systems are performed, including comparisons with experimental or atomistic simulation data. Here most of the systems are considered for validation. For a complete overview of the parametrization strategy used, see the Supplementary Notes. The Supplementary Notes provide details of the specific systems and methods used in the tests performed to parametrize and validate the new Martini 3 Lennard-Jones parameters.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Force-field parameters and procedures (for example, tutorials) are publicly available at http://cgmartini.nl. Simulation data (for example, trajectories) supporting the results of this paper are available from the corresponding authors upon reasonable request.

Code availability

Martinize2 (for which the manuscript is in preparation) and Martinate codes used in this work are publicly available at https://github.com/marrink-lab/. For more detailed information, see Supplementary Codes.

References

67. López, C. A. et al. Martini coarse-grained force field: extension to carbohydrates. *J. Chem. Theory Comput.* **5**, 3195–3210 (2009).
68. Monticelli, L. et al. The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* **4**, 819–834 (2008).
69. Grunewald, F., Rossi, G., de Vries, A. H., Marrink, S. J. & Monticelli, L. Transferable MARTINI model of poly(ethylene oxide). *J. Phys. Chem. B* **122**, 7436–7449 (2018).
70. de Jong, D. H. et al. Improved parameters for the martini coarse-grained protein force field. *J. Chem. Theory Comput.* **9**, 687–97 (2013).
71. Herzog, F. A., Braun, L., Schoen, I. & Vogel, V. Improved side chain dynamics in MARTINI simulations of protein–lipid interfaces. *J. Chem. Theory Comput.* **12**, 2446–2458 (2016).
72. Poma, A. B., Cieplak, M. & Theodorakis, P. E. Combining the MARTINI and structure-based coarse-grained approaches for the molecular dynamics studies of conformational transitions in proteins. *J. Chem. Theory Comput.* **13**, 1366–1374 (2017).
73. Periole, X., Cavalli, M., Marrink, S.-J. & Ceruso, M. A. Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J. Chem. Theory Comput.* **5**, 2531–2543 (2009).
74. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. & Marrink, S. J. Computational lipidomics with insane: a versatile tool for generating Custom membranes for molecular simulations. *J. Chem. Theory Comput.* **11**, 2144–2155 (2015).
75. Melo, M. N., Ingólfsson, H. I. & Marrink, S. J. Parameters for Martini sterols and hopanoids based on a virtual-site description. *J. Chem. Phys.* **143**, 243152 (2015).
76. López, C. A., Sovova, Z., van Eerden, F. J., de Vries, A. H. & Marrink, S. J. Martini force field parameters for glycolipids. *J. Chem. Theory Comput.* **9**, 1694–1708 (2013).
77. 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).
78. de Jong, D. H., Baoukina, S., Ingólfsson, H. I. & Marrink, S. J. Martini straight: boosting performance using a shorter cutoff and GPUs. *Comput. Phys. Commun.* **199**, 1–7 (2016).
79. Hockney, R. W., Goel, S. P. & Eastwood, J. W. Quiet high-resolution computer models of a plasma. *J. Comput. Phys.* **14**, 148–158 (1974).
80. Páll, S. & Hess, B. A flexible algorithm for calculating pair interactions on SIMD architectures. *Comput. Phys. Commun.* **184**, 2641–2650 (2013).
81. Verlet, L. Computer 'experiments' on classical fluids. I. Thermodynamical properties of Lennard–Jones molecules. *Phys. Rev.* **159**, 98–103 (1967).
82. Tironi, I. G., Sperb, R., Smith, P. E. & Van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **102**, 5451–5459 (1995).
83. Essmann, U. et al. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
84. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
85. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
86. Abraham, M. J. et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
87. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
88. Wassenaar, T. A., Ingólfsson, H. I., Prieß, M., Marrink, S. J. & Schäfer, L. V. Mixing MARTINI: electrostatic coupling in hybrid atomistic-coarse-grained biomolecular simulations. *J. Phys. Chem. B.* **117**, 3516–3530 (2013).
89. Wassenaar, T. A. et al. High-throughput simulations of dimer and trimer assembly of membrane proteins. The DAFT approach. *J. Chem. Theory Comput.* **11**, 2278–91 (2015).
90. Humphrey, W., Dalke, A. & Schulten, K. VMD—visual molecular dynamics. *J. Molec. Graph.* **14**, 33–38 (1996).
91. Gowers, R. J. et al. MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. in *Proc. 15th Python Sci. Conference* 98–105 (2016).

Acknowledgements

We thank all members of the S.J.M. group and also external users for testing Martini 3 in its open-beta version. In particular, we thank C. F. E. Schroer, P. W. J. M. Frederix, W. Pezeshkian, M. N. Melo, H. I. Ingólfsson, M. Tsanai, M. König, P. A. Vainikka, T. Zijp, L. Gaifas, J. H. van der Woude, M. Espinoza Cangahuala, M. Scharte, J. Cruiming, L. M. van der Sleen, V. Verduijn, A. H. Beck Frederiksen, B. Schiøtt, M. Sikora, P. Schmalhorst, R. A. Moreira, A. B. Poma, K. Pluhackova, C. Arnarez, C. A. López, E. Jefferys and M. S. P. Sansom for their preliminary tests with a lot of different systems including aedamers, sugars, amino acids, deep eutectic solvents, lipids, peptides and proteins. We also thank the Center for Information Technology of the University of Groningen for providing access to the Peregrine high-performance computing cluster. We acknowledge the National Computing Facilities Foundation of The Netherlands Organization for Scientific Research (NWO), CSC–IT Center for Science Ltd (Espoo, Finland) and CINES (France) for providing computing time. Work in the S.J.M. group was supported by an European Research Council advanced grant no. 'COMP-MICR-CROW-MEM'. R.A. thanks the NWO (Graduate Programme Advanced Materials, no. 022.005.006) for financial support. L.M. acknowledges the Institut National de la Santé et de la Recherche Medicale and the Agence Nationale de la Recherche for funding (grant no. ANR-17-CE11-0003) and GENCI-CINES for computing time (grant no. A0060710138). S.T. acknowledges the support from the European Commission via a Marie Skłodowska-Curie Actions individual fellowship (MicroMod-PSII, grant agreement no. 748895). M.J. thanks the Emil Aaltonen foundation for financial support. I.V. thanks the Academy of Finland (Center of Excellence program (grant no. 307415)), Sigrid Juselius Foundation, the Helsinki Institute of Life Science fellow program and the HFSP (research grant no. RGP0059/2019). R.B.B. and J.D. were supported by the intramural research program of the NIDDK, NIH. Their work used the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). H.M.-S. acknowledges the Czech Science Foundation (grant no. 19-19561S). J.B. acknowledges funding from the TOP grant from S.J.M. (NWO) and the EPSRC program grant no. EP/P021123/1. Work in D.P.T.'s group is supported by the Natural Sciences and Engineering Research Council (Canada) and Compute Canada, funded by the Canada Foundation for Innovation. D.P.T. acknowledges further support from the Canada Research Chairs program. N.R. acknowledges funding from the Norwegian Research Council (FRIMEDBIO nos. 251247 and 288008) and computational resources from UNINETT SIGMA2 AS (grant no. NN4700K). H.M.K. acknowledges funding from the University of Calgary through the 'Eyes High Postdoctoral Fellowship' program.

Author contributions

P.C.T.S. and S.J.M. conceived the project with suggestions from R.A., A.H.V., J.B. and S.T. P.C.T.S. generated and optimized all bead parameters. P.C.T.S., R.A. and J.B. generated the topology and bonded parameters of all CG models with suggestions from S.T. and I.F. P.C.T.S., R.A., A.H.V. and F.G. performed the simulations and analysis involving transfer free energies, solvent and polymer properties. P.C.T.S., S.T., J.B. and J.M. performed the simulations and analysis involving lipid bilayers. P.C.T.S., I.F. and R.A. performed the simulations and analysis involving nucleobases. P.C.T.S., I.P. and A.H.V. generated the models and performed the simulations and analysis involving aedamers. P.C.T.S. and F.G. generated the models and performed the simulations and analysis involving ionic liquids and ionic water solutions. R.A. generated the models and performed the simulations and analysis involving bulk heterojunctions, with suggestions from L.M. regarding the fullerene model. P.C.T.S., J.B., H.A., R.A., B.M.H.B., S.T., J.M., V.N., X.P., M.J., H.M.K., J.D., V.C. and H.M.-S. performed the simulations and analysis involving amino acids, peptides and proteins. J.B., T.A.W., P.C.K. and S.T. developed some tools and scripts used to generate the CG models and to run the molecular dynamics simulations. L.M., R.B.B., P.T., N.R., I.V., A.H.V. and S.J.M. provided guidance and supervision in the studies performed by their respective group members and collaborators. P.C.T.S. and S.J.M. wrote the main manuscript, with contributions from all the authors. P.C.T.S. prepared the figures with contributions from R.A., B.M.H.B., H.M.K. and A.H.V. P.C.T.S. wrote the Methods with contributions from all the authors. P.C.T.S. wrote the Supplementary Information, with contributions from all the authors. All the authors revised and approved the final version of the manuscript, Methods and Supplementary Information.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41592-021-01098-3.

Correspondence and requests for materials should be addressed to P.C.T.S. or S.J.M.

Peer review information *Nature Methods* thanks the anonymous reviewers for their contribution for the peer review of this work. Arunima Singh was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

nature research

Corresponding author(s): Paulo Cesar Telles de Souza and Siewert-Jan Marrink.

Last updated by author(s): Jan 25, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Figures were compiled using VMD 1.9.4a12 and Inkscape 1.1. The graphs were plotted using xmgrace (5.1.25) and gnuplot (5.2)
Data analysis	All the analysis were performed using gmx analysis version (Gromacs 2018), VMD 1.9.4a12, xmgrace (5.1.25) and MDAnalysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Force-field parameters and procedures (e.g. tutorials) will be publicly available at http://cgmartini.nl before publication. All the simulation data are available upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The number of systems and simulation sampling is enough for analysis performed in this proof of concept study.
Data exclusions	No data has been excluded.
Replication	The methodology is dependent of the overall sampling that is determined by the number of simulations and by how long are trajectories.
Randomization	Does not apply to MD simulations.
Blinding	Does not apply to MD simulations.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).
Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.
These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.
We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.
While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.
These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.
To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.
Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.
If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com