



HAL
open science

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

Paulo Souza, Riccardo Alessandri, Jonathan Barnoud, Sebastian Thallmair,
Ignacio Faustino, Fabian Grünewald, Ilias Patmanidis, Haleh Abdizadeh, Bart
Bruininks, Tsjerk Wassenaar, et al.

► **To cite this version:**

Paulo Souza, Riccardo Alessandri, Jonathan Barnoud, Sebastian Thallmair, Ignacio Faustino, et al..
Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nature Methods*, 2021,
18 (4), pp.382-388. 10.1038/s41592-021-01098-3 . hal-03863208

HAL Id: hal-03863208

<https://cnrs.hal.science/hal-03863208>

Submitted on 24 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Martini 3: A General Purpose Force Field for Coarse-Grained Molecular Dynamics

Paulo C. T. Souza^{*1,2}, Riccardo Alessandri¹, Jonathan Barnoud^{1,3}, Sebastian Thallmair^{1,4}, Ignacio Faustino¹, Fabian Grunewald¹, Ilias Patmanidis¹, Haleh Abdizadeh¹, Bart M.H. Bruininks¹, Tsjerk A. Wassenaar¹, Peter C. Kroon¹, Josef Melcr¹, Vincent Nieto², Valentina Corradi⁵, Hanif M. Khan^{5,6}, Jan Domański^{7,8}, Matti Javanainen^{9,10}, Hector Martinez-Seara⁹, Nathalie Reuter⁶, Robert B. Best⁸, Ilpo Vattulainen^{10,11}, Luca Monticelli², Xavier Periole^{12,13}, D. Peter Tieleman⁵, Alex H. de Vries¹, Siewert J. Marrink^{*1}

¹ Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Material, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands.

² Molecular Microbiology and Structural Biochemistry, UMR 5086 CNRS & University of Lyon, 7 Passage du Vercors, F-69367, Lyon, France.

³ Intangible Realities Laboratory, University of Bristol, School of Chemistry, Cantock's Close, Bristol BS8 1TS, United Kingdom.

⁴ Frankfurt Institute for Advanced Studies, Ruth-Moufang-Str. 1, 60438 Frankfurt am Main, Germany.

⁵ Centre for Molecular Simulation and Department of Biological Sciences, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada.

⁶ Department of Chemistry and Computational Biology Unit, University of Bergen, N-5020 Bergen, Norway.

⁷ Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, United Kingdom.

⁸ Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States.

⁹ Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, CZ-16610 Prague 6, Czech Republic.

¹⁰ Computational Physics Laboratory, Tampere University, FI-33100 Tampere, Finland.

¹¹ Department of Physics, University of Helsinki, FI-00014 Helsinki, Finland.

¹² Department of Chemistry, Aarhus University, Aarhus C, Denmark.

¹³ School of Biological Sciences, The University of Auckland, Symonds Street 3A, Auckland, New Zealand.

* Corresponding authors: Paulo C. T. Souza – paulocts@gmail.com
Siewert J. Marrink – s.j.marrink@rug.nl

54 **ABSTRACT**

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

65

66 **INTRODUCTION**

67 The molecular dynamics (MD) simulation technique has become an
68 indispensable tool in natural sciences, offering a spatio-temporal resolution
69 unmatched by any experimental technique¹. A major bottleneck of MD is the
70 limited time and length scales that are accessible. To overcome this limitation,
71 coarse-grained (CG) models representing groups of atoms by effective beads,
72 have achieved widespread use². The Martini model is among the most
73 popular CG models in the field of biomolecular simulation, due to its easy-to-
74 use building block principle. Martini relies on a four-to-one mapping scheme
75 (i.e., on average four heavy atoms and associated hydrogens are mapped
76 into one CG bead), and has been parametrized using a top-down approach
77 with thermodynamic partitioning data as the main target^{3,4}. Non-bonded
78 interactions between neutral beads of Martini are solely described by

79 Lennard-Jones (LJ) potentials, while charged beads also include Coulombic
80 interactions. The interaction strength of the LJ potential (i.e. its well depth) is
81 used to discriminate between different levels of polarity of the CG beads. The
82 model features four main classes of CG bead types, denoted C, N, P, and Q
83 representing non-polar, intermediately polar, polar, and charged chemical
84 groups, respectively⁴. Sub-labels are used to make a further distinction within
85 a class in terms of degree of polarity or hydrogen donor/acceptor capabilities.
86 In principle, all beads have the same size, denoted as regular (R) beads. By
87 way of exception, special small (S) beads were introduced to model ring-like
88 compounds for which a four-to-one mapping scheme is inadequate⁴. To
89 reproduce correct stacking and hydrogen bonding distances between
90 nucleotides, even smaller tiny (T) beads were found necessary⁵.
91 Parametrization of the cross-interactions between S- and T-beads with R-
92 beads, however, was done on an *ad-hoc* basis.

93 The Martini force field is used in a wide range of applications in diverse fields
94 including structural biology⁶⁻⁸, biophysics^{9,10}, biomedicine¹¹,
95 nanotechnology^{12,13}, and materials design^{14,15}. With its growing use, however,
96 a number of shortcomings of the Martini model have recently been identified.
97 One of the most important problems is the observation that certain molecules
98 tend to interact too strongly. This has been reported for proteins and
99 carbohydrates in solution, as well as for membrane embedded proteins¹⁶⁻¹⁸.
100 The origin lies among others in small but systematic deviations in packing and
101 intermolecular interactions¹⁹. Besides, the coverage of chemical space for
102 broader applications was uneven, and in some cases, such as selectivity of
103 nucleobase pairing^{5,20}, consistency was difficult to obtain given the limited

104 bead types and sizes. To alleviate these problems, we undertook a re-
105 balancing of all non-bonded interaction terms of the Martini model, including
106 the addition of new beads and labels. The new version, called Martini 3,
107 enables more accurate simulations of molecular systems in general. In this
108 paper, we present the key features of Martini 3 combined with examples of
109 new applications and improvements in relation to the previous Martini model.

110

111 **RESULTS**

112 **Re-parametrization of the beads.** In Martini 3, the new parametrization
113 strategy was based on the construction of prototype models of polar and non-
114 polar molecules in all three Martini resolutions. Self- (R-R, S-S, and T-T) and
115 cross-interactions (R-S, R-T, and S-T) of the different bead sizes were
116 optimized to be well-balanced (**Supplementary Notes, sections B1 and B2**).
117 In terms of chemical types, the beads were separated into three blocks:
118 organic, ions, and water (**Supplementary Notes, section A1**). The organic
119 (containing P-, N-, and C-beads) and ion (Q-beads) blocks have been
120 subjected to independent parametrizations, where different trends in self-
121 interaction, solvation, and transfer free energy upon bead size change were
122 included (**Supplementary Notes, sections A1, B2, and B3**). In contrast to
123 the previous version, water is defined as a separate bead type (called W),
124 which enables optimization of water properties independently from other
125 targets; for example, the freezing of water at room temperature (a problem
126 sometimes encountered with the previous water model) no longer occurs. In
127 addition, it is available in three different sizes as well (**Supplementary Notes,**
128 **section B4**). Together with this optimization strategy, the new Martini 3 model

129 also features a fully revised interaction matrix (**Supplementary Notes,**
130 **section A2**) and new intermediate interaction levels, added to smoothen the
131 transition between chemical types (**Supplementary Notes, section A3**).
132 Bead assignment and validation of the models were not only based on
133 experimental transfer free energies, but also included solvent miscibility data
134 (**Supplementary Notes, sections A6 and C2, and Supplementary Results,**
135 **sections E4-E6**) and a series of benchmark tests, ranging from structural
136 properties of bilayers to dimerization potentials of mean force (PMFs) of
137 proteins (**Methods, section 3, Supplementary Notes, section B5 and**
138 **Supplementary Results, sections F1-F5**).

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

150 The accuracy of CG models containing ring or branched fragments, which rely
151 heavily on smaller bead types, is also greatly increased in Martini 3. For
152 example, the binary mixing behavior of various solvents (**Supplementary**
153 **Results, sections E4-E6**) and the transfer free energies of many linear,

154 branched, and ring-like compounds (**Fig 1C** and **Supplementary Results,**
155 **section E1**) are now in very good agreement with experimental data. The
156 mean absolute error of transfer free energies compared to the experimental
157 data is 2 kJ/mol, with 86% of the molecules presenting errors lower than 3
158 kJ/mol.

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

171 The proper balancing of R-, S-, and T-beads in Martini 3 also implies that the
172 mapping choice of an arbitrary molecule to its Martini representation is now
173 better defined. S- and T-beads are not only suited to represent ring-like
174 compounds, but also used for cases involving 3-to-1 and 2-to-1 mapping of
175 linear and branched chemical groups (**Supplementary Notes, section C1**).

176

177 **Covering the chemical space with new beads and labels.** Together with a
178 thorough revision of the interaction strengths, in Martini 3, we extend the
179 number of chemical bead types and the ability to modify the bead properties
180 depending on the chemical details of the underlying moieties. Each P-, N-,
181 and C-class now has 6 bead types with different degrees of polarity, which
182 enables a more precise definition of different chemical groups by assigning
183 them to certain bead types. Additionally, we introduce a new X-class of beads
184 to model halo-compounds (**Supplementary Notes, section A1**). In the
185 previous version of Martini, some of the bead types were already sub-
186 classified according to their ability to act as hydrogen bond donor, acceptor, or
187 both. This property can be now attributed to all bead types of intermediate or
188 polar nature (N- or P- class). The effective interaction strength between donor
189 and acceptor pairs is increased, whereas donor-donor and acceptor-acceptor
190 pairs are weakened (**Supplementary Notes, section A4**). For example,
191 Martini 3 correctly reproduces the trends in hydrogen bond-based pairing of
192 nucleobases^{5,20} without the use of special-purpose beads specifically for
193 nucleobases (right panel of **Fig. 1D**). Note that chemical groups that can act
194 as both donor and acceptor at the same time are always represented by the
195 pure beads of the P- and N-class in Martini 3.

196 Next to the fine-tuning based on hydrogen bonding capabilities, we introduce
197 the possibility to change the interactions based on the electronic polarizability.
198 Depending on inductive or conjugate effects caused by chemical
199 functionalization, non-polar molecules can be polarized, i.e., they can acquire
200 an electron-donor (or “enriched”, label “e”) or electron-acceptor (or “vacancy”,
201 label “v”) character, which can promote preferential interactions. Polarizable

202 groups in Martini 3 can be distinguished through the label “e/v” which can only
203 be applied to the C- and X-class. A nice example of their application is the
204 strong and specific interaction between electron donor and electron acceptor
205 aromatic rings in aedamers, a class of molecules that have been studied
206 extensively in the context of biomimetic folding and self-assembly^{23,24}. The
207 use of “e/v” allows Martini 3 to capture the preferential interaction between
208 1,5-dialkoxynaphthalene (DAN) and naphthalene diimide (NDI) (left panel of
209 **Fig. 2A**) experimentally observed via NMR titration²³ and atomistic simulation
210 data. Self-assembly of amide-linked tetramers shows preferential formation of
211 alternating stacks of DAN and NDI, which is also measured by NMR and
212 isothermal titration calorimetry investigations²⁴. On top of hydrogen bonding
213 and electron polarization labels, all beads can have their self-interaction fine-
214 tuned by other sub-labels (as further described in the **Supplementary Notes,**
215 **section A4**).

216 Chemical groups carrying monovalent charges +1/-1 are represented in
217 Martini by the class of Q-beads (**Supplementary Notes, section A1**). The
218 original Martini model only considers monovalent ions, and was solely
219 optimized for regular bead sizes which represented small ions and their first
220 hydration shell. In Martini 3, charged groups can have either R-, S-, or T-size.
221 The tiny size category allows modeling of small, bare ions, enabling
222 applications that involve ion binding where (part of) the hydration shell is lost.
223 This feature is exemplified by the binding of sodium ions (represented by a
224 charged tiny bead) to a buried small cavity localized in the core of the
225 adenosine A_{2A} receptor (**Fig. 2B**). X-ray crystallographic²⁵ and ligand binding
226 assays²⁶ confirm the importance of sodium ions for the structure and for the

227 allosteric modulation of the A_{2A} receptor. Note that an extensive validation of
228 the lipid models in Martini 3 was performed to allow simulations of
229 transmembrane and peripheral membrane proteins (see **Supplementary**
230 **Results, section F1**).

231 In addition to the smaller sizes, the Q-class was also expanded to five bead
232 types, following the classical Hofmeister series trend^{27,28} (**Supplementary**
233 **Notes, section B3 and Supplementary Results, section F2**). At one
234 extreme, the Q5 bead may be used to represent hard monovalent ions with
235 the smallest polarizability, e.g. inorganic ions such as R₂PO₄⁻. At the other
236 end of the Martini Hofmeister series, the Q1 type models polarizable soft
237 monovalent ions, like N(CH₃)₄⁺, and implicitly includes in the LJ potential ion-π
238 interactions. Such differences in behavior of the different Q-bead types are
239 exemplified by MD simulations of the anion transfer between aqueous
240 solutions and organophosphonium-based ionic liquids (**Fig. 2C and**
241 **Supplementary Results, section F2**). Harder ions such as Cl⁻ (modeled as
242 TQ5 with -1 charge) tend to stay in the water phase, together with Na⁺ ions
243 (TQ5⁺ bead). In contrast, softer ions like ClO₄⁻ (Q2⁻ bead) can exchange with
244 Br⁻ (SQ4⁻ bead) or (partially) PF₆⁻ (Q1⁻ bead) from the ionic liquid phase. In the
245 case of the biphasic system using trihexyltetradecylphosphonium bromide
246 ([P₆₆₆₁₄][Br]), direct comparison to experimental data shows good agreement
247 for the anion transfer trends^{28,29}. The new Q-bead types also impact
248 biologically relevant systems, as exemplified by preferential cation-π
249 interaction between choline groups (Q1⁺ bead) of phosphatidylcholine lipids
250 and aromatic residues of the *Bacillus thuringiensis* phosphatidylinositol-
251 specific phospholipase C (*BtPI-PLC*). In the previous version, such specific

252 interaction between soft ions and aromatic molecules were solely included in
253 the recently updated polarizable Martini implementation³⁰. However, in Martini
254 3, the different Q-bead types allow easier (implicit) incorporation of such
255 interactions without the need for additional partial charges.

256 On top of the new chemical types, all Q-beads can use the hydrogen-bonding
257 labels (called in this case “p/n”). They represent organic charged molecules or
258 fragments, such as R-CH₂-COO⁻ and R-CH₂-NH₃⁺, and also introduce
259 modifications in the Hofmeister trends of the pure Q-beads (**Supplementary**
260 **Notes, section A4**). Positively charged hydrogen donors (“p” label) interact
261 more strongly with non-polar beads, as expected in cation-π interactions. On
262 the other hand, negatively charged hydrogen acceptors (“n” label), have
263 stronger interactions with neutral polar beads, mimicking the stronger
264 hydrogen-bonds with anions. To complete the ion block, we explicitly include
265 a new D-bead for divalent ions (such as Ca²⁺), which are typically hard ions.

266

267 **Improving packing and protein-protein interactions.** Another change in
268 philosophy with respect to the previous Martini models is the definition of
269 bonded interactions. Instead of using the center of mass of the mapped
270 chemical groups to define the geometry of the molecule, we now use a size-
271 shape concept aimed at preserving the volume of molecules in comparison to
272 all-atom reference structures. This choice and the proper use of Martini 3
273 bead sizes (**Supplementary Notes, sections C1 and C2**) lead to more
274 realistic molecular packing. As a consequence, the hydration of protein
275 pockets and channels is improved, as illustrated by the Fragaceatoxin C
276 (FraC) nanopore inserted in a lipid bilayer (**Fig. 3A**). The pore of FraC

277 remains open over the whole trajectory in Martini 3, as indicated by X-ray
278 crystallography³¹ and electro-osmotic flow assays³², while in Martini 2 it is
279 closed.

280 Another example of accurate packing is the stacking predictions of thiophene
281 derivatives in bulk heterojunction solar cells composed of poly(3-hexyl-
282 thiophene) (P3HT) and phenyl-C61-butyric acid methyl ester (PCBM) (**Fig.**
283 **3B**). The morphology of these organic solar cells is a determinant for high-
284 efficiency devices³³. The scattering profiles computed with Martini 3 show
285 improved agreement with Martini 3 in relation to P3HT lamellar (peak around
286 $q \approx 0.45 \text{ \AA}^{-1}$) and stacking ($q \approx 1.65 \text{ \AA}^{-1}$) experimental distances^{33,34}.

287 The use of bonds based on molecular volume and the appropriate choice of
288 chemical bead types, sizes, and labels also controls the interaction density of
289 the model, which has an important impact on the strength of collective
290 interactions between molecules¹⁹. In order to test to what extent the changes
291 in non-bonded and bonded interactions reduce the over-estimated
292 aggregation of proteins, we performed extensive simulations comprised of
293 solutions of soluble proteins as well as membrane embedded proteins. These
294 systems were simulated under conditions in which proteins do not aggregate
295 and, preferentially, stay as monomers. For soluble proteins (**Fig. 3C**),
296 qualitative tests were performed with the headpiece domain of chicken villin³⁵,
297 and the modified and mutated cellulose-binding domain from *Cellulomonas*
298 *fimi* (EXG-CBM), which is an example of a protein completely free of charged
299 side chains that can maintain solubility, stability, and function³⁶. Trends are
300 well-captured in Martini 3, with both proteins mainly staying as monomers in
301 pure water (with only counter-ions to neutralize the system in the case of

302 villin). The villin headpiece showed salting-in behavior (i.e. less aggregation)
303 under addition of 0.4M of NaCl, which was also observed for certain soluble
304 proteins at low ionic strengths³⁷. On the other hand, EXG-CBM only showed
305 salting-out behavior (i.e. more aggregation), which was expected based on
306 experimental data³⁶. In contrast, both proteins aggregate in Martini 2, forming
307 a single and stable aggregate during the simulation.

308 Poly-leucine (K₂-L₂₆-K₂) was selected to evaluate the aggregation propensity in
309 membranes. Experimental evidence with this transmembrane (TM) protein
310 model indicates a preference for the monomeric state in a bilayer
311 environment³⁸⁻⁴⁰. Both Martini 2 and 3 show that the hydrophobic mismatch
312 between TM length and membrane thickness can play a role in the
313 aggregation, with Martini 3 showing a higher percentage of the monomeric
314 state (**Fig. 3D**). To quantitatively evaluate the strength of protein-protein
315 interactions in a membrane environment, we also considered the dimerization
316 of four selected transmembrane (TM) helices: the TM domains of the receptor
317 tyrosine kinases EphA1 and ErbB1; the red blood cell protein glycophorin A
318 (GpA); as well as the well-known model peptide WALP23 (left panel of **Fig.**
319 **3E**). For EphA1 and ErbB1, experimental dimerization free energies in a
320 membrane environment have been estimated using Förster resonance energy
321 transfer (FRET)^{41,42}. For GpA, dimerization free energies range from around -
322 15 kJ/mol (in various cell membrane environments)^{43,44} to -31.5 kJ/mol
323 (GALLEX assay in *E. coli* inner membranes)^{45,46} or -50.6 kJ/mol (steric trap in
324 POPC bilayers)⁴⁷. WALP peptides have been characterized thoroughly during
325 the past two decades, including their self-association⁴⁸. For each one of the
326 four peptide dimers, we compared experimental dimerization free energies

327 with the free energies predicted by the Martini 2 and Martini 3 model
328 predictions. Martini 3 shows not only to be able to capture the correct trends,
329 but also to quantitatively reproduce the experimental affinities. The binding
330 mode also becomes improved as highlighted for GpA (right panel of **Fig. 3E**).
331 The GpA homodimer structure with Martini 3 closely resembles experimental
332 results obtained with NMR spectroscopy and crystallography⁴⁹⁻⁵¹.

333 In summary, for both soluble and transmembrane proteins, we observed that
334 the Martini 3 models are in much better agreement with experimental data
335 than before. Another advantage of the current Martini 3 protein model is that
336 the default bead type representing the protein backbone in Martini 3 (a regular
337 P2 bead) no longer depends on the secondary structure. In addition, the
338 representation of protein flexibility can now be improved by the use of Gō-like
339 models⁵².

340

341 **DISCUSSION**

342 In this paper we have described the new version of the Martini force field,
343 which shows numerous improvements in relation to the previous version.
344 However, inherent limitations to the process of coarse-graining, related to
345 transferability and representability problems⁵³⁻⁵⁵ are still part of the model. An
346 important drawback is the limited structural detail, that is a consequence of
347 representing multiple atoms with isotropic interaction sites. This is most
348 noticeable for the Martini water model, which represents four water molecules
349 with a single LJ site and will certainly not capture any of the higher order
350 structural correlations of real water. The role of explicit water in a CG model

351 such as Martini is mostly to provide a good solvent for polar compounds
352 resulting in realistic partitioning. For applications requiring finer details,
353 structure-based CG models are more suitable^{56,57}. Another fundamental
354 limitation is the entropy-enthalpy compensation. The loss of internal degrees
355 of freedom for groups of atoms represented by a CG bead inevitably reduces
356 the entropy of the system. Since the Martini force field is based on
357 reproducing free energies, this requires a concomitant reduction in the
358 enthalpy. As consequence, inaccurate entropy-enthalpy balance affects the
359 temperature dependence of several properties and reduces the transferability
360 to different state points. To probe transferability, we performed temperature
361 dependent calculations for a number of solvent systems as well as lipid
362 membranes (**Supplementary Results F**). Temperature dependent properties,
363 like the heat expansion coefficient and heat capacity of water and n-octane,
364 are very well captured by Martini 3, but this is not true for the hydrophobic
365 effect, that shows the opposite trend with respect to atomistic models, in line
366 with previous findings⁵⁸. Note that the use of higher-resolution S- or T-
367 particles does not remedy this problem, as these bead types were
368 parameterized mainly to be compatible with the regular (R-type) beads and
369 should be used primarily to represent parts of the system that cannot be
370 adequately mapped with R-particles. Potential improvements with respect to
371 the temperature transferability of our model could make use of environment
372 dependent potentials⁵⁹ or CG beads with embedded sites, such as the
373 polarizable water models^{60,61}, where incorporation of quadrupole moment
374 might be required⁵⁸. Bottom-up CG models that are derived with minimization
375 of the information loss⁵⁴ as parameterization target might also perform better.

376 For a more in depth discussion of these and related issues with respect to the
377 Martini coarse-graining philosophy, we refer to previous papers^{62,63}.
378 Keeping these limitations in mind, Martini 3 offers a versatile and easy to use
379 generic force field to simulate molecular processes at a semi-quantitative level
380 of accuracy. In relation to the previous model, the excessive over-estimated
381 aggregation¹⁹ is substantially reduced. We expect that Martini 3 will allow for
382 more realistic predictions of protein interactions, as well as more accurate
383 simulations of molecular systems in general. The increased number of bead
384 types and interaction levels makes the model even more versatile, covering a
385 larger part of chemical space with appropriate building blocks. Based on this
386 new foundation, further optimizations for different classes of molecules are
387 currently ongoing, including the use of Gō-potentials to alleviate limitations of
388 protein conformational flexibility, a re-optimization of the bonded potentials of
389 lipids and other biomolecular classes, as well as a compatible polarizable
390 water model for applications requiring more realistic screening of electrostatic
391 interactions. Finally, we foresee new application horizons for the Martini
392 model in the field of materials science^{64,65} and high-throughput drug design⁶⁶.

393

394 **ONLINE CONTENT**

395 Methods, including statements of data and code availability, supplementary
396 information, and any other associated content and references, are available in
397 the online version of the paper at <https://doi.org/10.1038/sXXXXX-XXX-XXXX->
398 X.

399

400 **ACKNOWLEDGMENTS**

401 We thank all members of the S.J.M. group and also external users for testing
402 Martini 3 in its open-beta version. In particular, we would like to thank C.F.E.
403 Schroer, P.W.J.M. Frederix, W. Pezeshkian, M.N. Melo, H.I. Ingólfsson, M.
404 Tsanai, M. König, P. A. Vainikka, T. Zijp, L. Gaifas, J.H. van der Woude, M.
405 Espinoza Cangahuala, M. Scharte, J. Cruiming, L.M. van der Sleen, V.
406 Verduijn, A.H. Beck Frederiksen, B. Schiøtt, M. Sikora, P. Schmalhorst, K.
407 Pluhackova, C. Arnarez, C.A. López, E. Jefferys, and M.S.P. Sansom for their
408 preliminary tests with a lot of different systems including aedamers, sugars,
409 amino acids, deep eutectic solvents, lipids, peptides, and proteins. We also
410 thank the Center for Information Technology of the University of Groningen for
411 providing access to the Peregrine high-performance computing cluster. We
412 acknowledge the National Computing Facilities Foundation (NCF) of The
413 Netherlands Organization for Scientific Research (NWO), CSC – IT Center for
414 Science Ltd (Espoo, Finland), and CINES (France) for providing computing
415 time. Work in the S.J.M. group was supported by an ERC Advanced Grant
416 “COMP-MICR-CROW-MEM”. R.A. thanks The Netherlands Organisation for
417 Scientific Research NWO (Graduate Programme Advanced Materials, No.
418 022.005.006) for financial support. L.M. acknowledges the Institut National de
419 la Santé et de la Recherche Medicale (INSERM) and the Agence Nationale
420 de la Recherche (ANR) for funding (grant no. ANR-17-CE11-0003) and
421 GENCI-CINES for computing time (grant no. A0060710138). S.T.
422 acknowledges the support from the European Commission via a Marie
423 Skłodowska-Curie Actions individual fellowship (MicroMod-PSII, grant
424 agreement 748895). M.J. thanks the Emil Aaltonen foundation for financial

425 support. I.V. thanks the Academy of Finland (Center of Excellence program
426 (grant no. 307415)), Sigrid Juselius Foundation, the Helsinki Institute of Life
427 Science fellow program, and the HFSP (research grant Ref.-No:
428 RGP0059/2019). R.B. and J.D. were supported by the intramural research
429 program of the NIDDK, NIH. Their work utilized the computational resources
430 of the NIH HPC Biowulf cluster. (<http://hpc.nih.gov>). H.M.-S. acknowledges
431 the Czech Science Foundation (19-19561S). J.B. acknowledges funding from
432 the TOP grant of S.J.M. (NWO) and the EPSRC program grant
433 EP/P021123/1. Work in D.P.T.'s group is supported by the Natural Sciences
434 and Engineering Research Council (Canada) and Compute Canada, funded
435 by the Canada Foundation for Innovation. D.P.T. acknowledges further
436 support from the Canada Research Chairs program.

437

438 **AUTHOR CONTRIBUTIONS**

439 P.C.T.S. and S.J.M. conceived the project with suggestions from R.A., A.H.V.,
440 J.B., and S.T.; P.C.T.S. generated and optimized all bead parameters;
441 P.C.T.S., R.A., and J.B. generated the topology and bonded parameters of all
442 CG models with suggestions from S.T. and I.F.; P.C.T.S., R.A., A.H.V., and
443 F.G. performed the simulations and analysis involving transfer free energies,
444 solvent and polymer properties; P.C.T.S., S.T., J.B., and J.M. performed the
445 simulations and analysis involving lipid bilayers; P.C.T.S., I.F., and R.A.
446 performed the simulations and analysis involving nucleobases; P.C.T.S., I.P.,
447 and A.H.V. generated the models, performed the simulations and analysis
448 involving aedamers; P.C.T.S., and F.G. generated the models, performed the
449 simulations and analysis involving ionic liquids and ionic water solutions; R.A.

450 generated the models, performed the simulations and analysis involving bulk
451 heterojunctions, with suggestions from L.M. regarding the fullerene model;
452 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.,
453 V.C., and H.M.-S. performed the simulations and analysis involving amino
454 acids, peptides and proteins; J.B., T.W., P.K., and S.T. developed some tools
455 and scripts used to generate the CG models and to run the MD simulations;
456 L.M., R.B., P.T., N.R., I.V., A.H.V., and S.J.M. provided guidance and
457 supervision in the studies performed by their respective group members and
458 collaborators; P.C.T.S. and S.J.M. wrote the main manuscript, with
459 contributions from all the authors; P.C.T.S. prepared the figures with
460 contributions from R.A., B.M.H.B., H.M.K., and A.H.V.; P.C.T.S. wrote the
461 Methods with contributions from all the authors. P.C.T.S. wrote the
462 Supplementary Information, with contributions from all the authors. All the
463 authors revised and approved the final version of the manuscript, Online
464 Methods, and Supplementary Information.

465

466 **COMPETING INTERESTS**

467 The authors declare no competing interests.

468

469 **REFERENCES**

- 470 1. Bottaro, S. & Lindorff-Larsen, K. Biophysical experiments and biomolecular
471 simulations: A perfect match? *Science* **361**, 355–360 (2018).
- 472 2. Ingólfsson, H. I. *et al.* The power of coarse graining in biomolecular
473 simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4**, 225–248 (2014).

- 474 3. Marrink, S. J., De Vries, A. H. & Mark, A. E. Coarse Grained Model for
475 Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **108**, 750–760 (2004).
- 476 4. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A.
477 H. The MARTINI force field: coarse grained model for biomolecular
478 simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
- 479 5. Uusitalo, J. J., Ingólfsson, H. I., Akhshi, P., Tieleman, D. P. & Marrink, S. J.
480 Martini Coarse-Grained Force Field: Extension to DNA. *J. Chem. Theory*
481 *Comput.* **11**, 3932–3945 (2015).
- 482 6. Abellón-Ruiz, J. *et al.* Structural basis for maintenance of bacterial outer
483 membrane lipid asymmetry. *Nat. Microbiol.* **2**, 1616–1623 (2017).
- 484 7. Yen, H. Y. *et al.* PtdIns(4,5)P₂ stabilizes active states of GPCRs and
485 enhances selectivity of G-protein coupling. *Nature* **559**, 423–427 (2018).
- 486 8. Van Eerden, F. J., Melo, M. N., Frederix, P. W. J. M., Periole, X. & Marrink,
487 S. J. Exchange pathways of plastoquinone and plastoquinol in the
488 photosystem II complex. *Nat. Commun.* **8**, 15214 (2017).
- 489 9. Vögele, M., Köfinger, J. & Hummer, G. Hydrodynamics of Diffusion in Lipid
490 Membrane Simulations. *Phys. Rev. Lett.* **120**, (2018).
- 491 10. Agostino, M. D., Risselada, H. J., Lürick, A., Ungermann, C. & Mayer, A. A
492 tethering complex drives the terminal stage of SNARE-dependent membrane
493 fusion. *Nature* **551**, 634–638 (2017).
- 494 11. Jeena, M. T. *et al.* Mitochondria localization induced self-assembly of
495 peptide amphiphiles for cellular dysfunction. *Nat. Commun.* **8**, 26 (2017).
- 496 12. Jiang, Z. *et al.* Subnanometre ligand-shell asymmetry leads to Janus-like
497 nanoparticle membranes. *Nat. Mater.* **14**, 912–917 (2015).

- 498 13. Maingi, V. *et al.* Stability and dynamics of membrane-spanning DNA
499 nanopores. *Nat. Commun.* **8**, 14784 (2017).
- 500 14. Frederix, P. W. J. M. *et al.* Exploring the sequence space for (tri-)peptide
501 self-assembly to design and discover new hydrogels. *Nat. Chem.* **7**, 30–37
502 (2015).
- 503 15. Bochicchio, D., Salvalaglio, M. & Pavan, G. M. Into the dynamics of a
504 supramolecular polymer at submolecular resolution. *Nat. Commun.* **8**, 147
505 (2017).
- 506 16. Stark, A. C., Andrews, C. T. & Elcock, A. H. Toward optimized potential
507 functions for protein-protein interactions in aqueous solutions: osmotic second
508 virial coefficient calculations using the MARTINI coarse-grained force field. *J.*
509 *Chem. Theory Comput.* **9**, 4176–4185 (2013).
- 510 17. Javanainen, M., Martinez-Seara, H. & Vattulainen, I. Excessive
511 aggregation of membrane proteins in the Martini model. *PLoS One* **12**,
512 e0187936 (2017).
- 513 18. Schmalhorst, P. S., Deluweit, F., Scherrers, R., Heisenberg, C.-P. &
514 Sikora, M. Overcoming the Limitations of the MARTINI Force Field in
515 Simulations of Polysaccharides. *J. Chem. Theory Comput.* **13**, 5039–5053
516 (2017).
- 517 19. Alessandri, R. *et al.* Pitfalls of the Martini Model. *J. Chem. Theory*
518 *Comput.* **15**, 5448–5460 (2019).
- 519 20. Uusitalo, J. J., Ingólfsson, H. I., Marrink, S. J. & Faustino, I. Martini
520 Coarse-Grained Force Field: Extension to RNA. *Biophys. J.* **113**, 246–256
521 (2017).

- 522 21. Ben-Naim, A. *Molecular theory of solutions*. (Oxford University Press,
523 2006).
- 524 22. Ploetz, E. A., Benteñitis, N. & Smith, P. E. Kirkwood-Buff integrals for ideal
525 solutions. *J. Chem. Phys.* **132**, 164501 (2010).
- 526 23. Zych, A. J. & Iverson, B. L. Synthesis and Conformational
527 Characterization of Tethered, Self-Complexing 1,5-
528 Dialkoxynaphthalene/1,4,5,8-Naphthalenetetracarboxylic Diimide Systems. *J.*
529 *Am. Chem. Soc.* **122**, 8898–8909 (2000).
- 530 24. Gabriel, G. J. & Iverson, B. L. Aromatic oligomers that form hetero
531 duplexes in aqueous solution. *J. Am. Chem. Soc.* **124**, 15174–15175 (2002).
- 532 25. Liu, W. *et al.* Structural basis for allosteric regulation of GPCRs by sodium
533 ions. *Science* **337**, 232–236 (2012).
- 534 26. Gao, Z. G. & Ijzerman, A. P. Allosteric modulation of A(2A) adenosine
535 receptors by amiloride analogues and sodium ions. *Biochem. Pharmacol.* **60**,
536 669–676 (2000).
- 537 27. Okur, H. I. *et al.* Beyond the Hofmeister Series: Ion-Specific Effects on
538 Proteins and Their Biological Functions. *J. Phys. Chem. B* **121**, 1997–2014
539 (2017).
- 540 28. Dupont, D., Depuydt, D. & Binnemans, K. Overview of the Effect of Salts
541 on Biphasic Ionic Liquid/Water Solvent Extraction Systems: Anion Exchange,
542 Mutual Solubility, and Thermomorphic Properties. *J. Phys. Chem. B* **119**,
543 6747–6757 (2015).
- 544 29. Naert, P., Rabaey, K. & Stevens, C. V. Ionic liquid ion exchange:
545 Exclusion from strong interactions condemns cations to the most weakly

546 interacting anions and dictates reaction equilibrium. *Green Chem.* **20**, 4277–
547 4286 (2018).

548 30. Khan, H. M. *et al.* Capturing Choline-Aromatics Cation- π Interactions in the
549 MARTINI Force Field. *J. Chem. Theory Comput.* **16**, 2550–2560 (2020).

550 31. Tanaka, K., Caaveiro, J. M. M., Morante, K., González-Manãs, J. M. &
551 Tsumoto, K. Structural basis for self-assembly of a cytolytic pore lined by
552 protein and lipid. *Nat. Commun.* **6**, 6337 (2015).

553 32. Huang, G., Willems, K., Soskine, M., Wloka, C. & Maglia, G. Electro-
554 osmotic capture and ionic discrimination of peptide and protein biomarkers
555 with FraC nanopores. *Nat. Commun.* **8**, 935 (2017).

556 33. Alessandri, R., Uusitalo, J. J., De Vries, A. H., Havenith, R. W. A. &
557 Marrink, S. J. Bulk Heterojunction Morphologies with Atomistic Resolution
558 from Coarse-Grain Solvent Evaporation Simulations. *J. Am. Chem. Soc.* **139**,
559 3697–3705 (2017).

560 34. Chiu, M. Y., Jeng, U. S., Su, C. H., Liang, K. S. & Wei, K. H. Simultaneous
561 use of small- and wide-angle X-ray techniques to analyze nanometerscale
562 phase separation in polymer heterojunction solar cells. *Adv. Mater.* **20**, 2573–
563 2578 (2008).

564 35. Petrov, D. & Zagrovic, B. Are Current Atomistic Force Fields Accurate
565 Enough to Study Proteins in Crowded Environments? *PLoS Comput. Biol.* **10**,
566 e1003638 (2014).

567 36. Højgaard, C. *et al.* A Soluble, Folded Protein without Charged Amino Acid
568 Residues. *Biochemistry* **55**, 3949–3956 (2016).

- 569 37. Ruckenstein, E. & Shulgin, I. L. Effect of salts and organic additives on the
570 solubility of proteins in aqueous solutions. *Advances in Colloid and Interface*
571 *Science* **123–126**, 97–103 (2006).
- 572 38. Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T. & Engelman, D. M.
573 Interhelical hydrogen bonding drives strong interactions in membrane
574 proteins. *Nat. Struct. Biol.* **7**, 154–160 (2000).
- 575 39. Zhou, F. X., Merianos, H. J., Brunger, A. T. & Engelman, D. M. Polar
576 residues drive association of polyleucine transmembrane helices. *Proc. Natl.*
577 *Acad. Sci. U. S. A.* **98**, 2250–2255 (2001).
- 578 40. Grau, B. *et al.* The role of hydrophobic matching on transmembrane helix
579 packing in cells. *Cell Stress* **1**, 90–106 (2017).
- 580 41. Chen, L., Merzlyakov, M., Cohen, T., Shai, Y. & Hristova, K. Energetics of
581 ErbB1 transmembrane domain dimerization in lipid bilayers. *Biophys. J.* **96**,
582 4622–4630 (2009).
- 583 42. Artemenko, E. O., Egorova, N. S., Arseniev, A. S. & Feofanov, A. V.
584 Transmembrane domain of EphA1 receptor forms dimers in membrane-like
585 environment. *Biochim. Biophys. Acta* **1778**, 2361–7 (2008).
- 586 43. Sarabipour, S. & Hristova, K. Glycophorin A transmembrane domain
587 dimerization in plasma membrane vesicles derived from CHO, HEK 293T, and
588 A431 cells. *Biochim. Biophys. Acta - Biomembr.* **1828**, 1829–1833 (2013).
- 589 44. Chen, L., Novicky, L., Merzlyakov, M., Hristov, T. & Hristova, K.
590 Measuring the Energetics of Membrane Protein Dimerization in Mammalian
591 Membranes. *J. Am. Chem. Soc.* **132**, 3628–3635 (2010).

592 45. Nash, A., Notman, R. & Dixon, A. M. De novo design of transmembrane
593 helix-helix interactions and measurement of stability in a biological membrane.
594 *Biochim. Biophys. Acta - Biomembr.* **1848**, 1248–1257 (2015).

595 46. Finger, C. *et al.* The Stability of Transmembrane Helix Interactions
596 Measured in a Biological Membrane. *J. Mol. Biol.* **358**, 1221–1228 (2006).

597 47. Hong, H., Blois, T. M., Cao, Z. & Bowie, J. U. Method to measure strong
598 protein-protein interactions in lipid bilayers using a steric trap. *Proc. Natl.*
599 *Acad. Sci. U. S. A.* **107**, 19802–19807 (2010).

600 48. Sparr, E. *et al.* Self-association of transmembrane α -helices in model
601 membranes: Importance of helix orientation and role of hydrophobic
602 mismatch. *J. Biol. Chem.* **280**, 39324–39331 (2005).

603 49. MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. Transmembrane
604 helix dimer: Structure and implications. *Science (80-.).* **276**, 131–133 (1997).

605 50. Trenker, R., Call, M. E. & Call, M. J. Crystal Structure of the Glycophorin A
606 Transmembrane Dimer in Lipidic Cubic Phase. *J. Am. Chem. Soc.* **137**,
607 15676–15679 (2015).

608 51. Domański, J., Sansom, M. S. P., Stansfeld, P. J. & Best, R. B. Balancing
609 Force Field Protein–Lipid Interactions To Capture Transmembrane Helix–
610 Helix Association. *J. Chem. Theory Comput.* **14**, 1706–1715 (2018).

611 52. Souza, P. C. T., Thallmair, S., Marrink, S. J. & Mera-Adasme, R. An
612 Allosteric Pathway in Copper, Zinc Superoxide Dismutase Unravels the
613 Molecular Mechanism of the G93A Amyotrophic Lateral Sclerosis-Linked
614 Mutation. *J. Phys. Chem. Lett.* **10**, 7740–7744 (2019).

615 53. Brini, E. *et al.* Systematic coarse-graining methods for soft matter
616 simulations-a review. *Soft Matter* **9**, 2108–2119 (2013).

617 54. Foley, T. T., Shell, M. S. & Noid, W. G. The impact of resolution upon
618 entropy and information in coarse-grained models. *J. Chem. Phys.* **143**,
619 243104 (2015).

620 55. Wagner, J. W., Dama, J. F., Durumeric, A. E. P. & Voth, G. A. On the
621 representability problem and the physical meaning of coarse-grained models.
622 *J. Chem. Phys.* **145**, 044108 (2016).

623 56. Wörner, S. J., Bereau, T., Kremer, K. & Rudzinski, J. F. Direct route to
624 reproducing pair distribution functions with coarse-grained models via
625 transformed atomistic cross correlations. *J. Chem. Phys.* **151**, 244110 (2019).

626 57. Noid, W. G., Chu, J. W., Ayton, G. S. & Voth, G. A. Multiscale coarse-
627 graining and structural correlations: Connections to liquid-state theory. *J.*
628 *Phys. Chem. B* **111**, 4116–4127 (2007).

629 58. Wu, Z., Cui, Q. & Yethiraj, A. Driving force for the association of
630 hydrophobic peptides: The importance of electrostatic interactions in coarse-
631 grained water models. *J. Phys. Chem. Lett.* **2**, 1794–1798 (2011).

632 59. Jin, J., Yu, A. & Voth, G. A. Temperature and Phase Transferable Bottom-
633 up Coarse-Grained Models. *J. Chem. Theory Comput.* (2020).
634 doi:10.1021/acs.jctc.0c00832

635 60. Yesylevskyy, S. O., Schäfer, L. V, Sengupta, D. & Marrink, S. J.
636 Polarizable water model for the coarse-grained MARTINI force field. *PLoS*
637 *Comput. Biol.* **6**, e1000810 (2010).

638 61. Michalowsky, J., Schäfer, L. V., Holm, C. & Smiatek, J. A refined
639 polarizable water model for the coarse-grained MARTINI force field with long-
640 range electrostatic interactions. *J. Chem. Phys.* **146**, 054501 (2017).

- 641 62. Marrink, S. J. & Tieleman, D. P. Perspective on the Martini model. *Chem.*
642 *Soc. Rev.* **42**, 6801–22 (2013).
- 643 63. Bruininks, B. M. H., Souza, P. C. T. & Marrink, S. J. A Practical View of
644 the Martini Force Field. in *Methods in Molecular Biology* **2022**, 105–127
645 (Humana Press Inc., 2019).
- 646 64. Liu, J. *et al.* Enhancing Molecular n-Type Doping of Donor-Acceptor
647 Copolymers by Tailoring Side Chains. *Adv. Mater.* **30**, 1704630 (2018).
- 648 65. Vazquez-Salazar, L. I., Selle, M., de Vries, A., Marrink, S. J. & Souza, P.
649 C. T. Martini coarse-grained models of imidazolium-based ionic liquids: from
650 nanostructural organization to liquid-liquid extraction. *Green Chem.* **22**, 7376–
651 7386 (2020).
- 652 66. Souza, P. C. T. *et al.* Protein–ligand binding with the coarse-grained
653 Martini model. *Nat. Commun.* **11**, 1–11 (2020).

654 **FIGURE LEGENDS**

655

656 **Figure 1: Rebalancing R-, S-, and T-beads** – (A) Snapshots of simulation boxes containing
657 mixtures of dodecane and water in three resolutions. (B) Radial distribution functions (g_{ij}) for
658 all bead combinations in the multi-resolution mixture of water. (C) Water-oil transfer free
659 energies (ΔG) computed for around 260 data points using Martini 3. (D) Hydrogen bonding
660 potential of mean force (PMF) between nucleobases. On the left, comparison between Martini
661 2 and 3 for the cytosine-guanine base pair. On the right, comparison of the cytosine-guanine
662 (C-G) and guanine-guanine (G-G) base pairs using Martini 3. In both plots, CHARMM and
663 AMBER atomistic data are also reported⁵ for comparison. Errors are estimated with
664 bootstrapping and displayed as transparent shades. In the case of Martini, errors are smaller
665 than 0.1 kJ/mol, and hence are not visible in the graphs.

666

667 **Figure 2: New chemical bead types, sub-labels, and applications** – (A) Self-assembly of
668 aedamers. The left panel shows the dimerization free energies (ΔG_{dim}) of pegylated
669 monomers of 1,5-dialkoxynaphthalene (DAN) and naphthalene diimide (NDI). Errors are
670 estimated with bootstrapping. The right panel shows the self-assembled duplex dimer formed
671 by amide-linked tetramers of NDI (green) and DAN (orange). (B) As indicated by X-ray
672 crystallography²⁵, sodium ions (charged TQ5 bead) can bind to a buried small cavity in the
673 core of the adenosine A_{2A} receptor. (C) Charged Q-beads in Martini 3 follow the classical
674 Hofmeister series, as exemplified by the anion transfer between salt aqueous solutions and
675 organophosphonium-based ionic liquids (right panel). Errors in the average anion transfer
676 percentage are estimated by block averaging. (D) Preferential cation- π interaction between
677 choline groups (Q1 bead) of phosphatidylcholine lipids and aromatic residues of the *Bacillus*
678 *thuringiensis* phosphatidylinositol-specific phospholipase C (*BtPI-PLC*). The depth of insertion
679 of each amino acid of *BtPI-PLC* is in very good agreement with the insertion obtained from an
680 atomistic MD simulation³⁰.

681

682 **Figure 3: Improving packing, cavities and reducing protein stickiness** – (A) Hydration of
683 Fragaceatoxin C (FraC) nanopore inserted in a lipid bilayer. (B) Scattering profiles and a
684 Martini 3 snapshot of a bulk heterojunction morphology of poly(3-hexyl-thiophene) (P3HT, in
685 red) and phenyl-C61-butyric acid methyl ester (PCBM, in blue) formed after solvent
686 evaporation and annealing simulations. $I(q)$ corresponds to scattering intensity and q is the
687 reciprocal space vector. (C) Aggregation levels of the soluble proteins villin headpiece and
688 the modified EXG-CBM in different salt concentrations. (D) Aggregation levels of polyileucine
689 helices in POPC and DLPC bilayers. Errors in the average monomer percentage of (C) and
690 (D) are estimated by block averaging. (E) Dimerization of transmembrane helices. Left panel
691 shows a comparison between experimental and calculated values for the mole fraction
692 standard Gibbs free energy of dimerization (ΔG_{ass}^X) of the following transmembrane protein
693 domains: ErbB1, EphA1, WALP23 and GpA. Simulation errors are estimated with
694 bootstrapping while experimental data was obtained in the literature⁴¹⁻⁴⁸. In the case of GpA,
695 error was estimated by the mean absolute error of four independent experimental data⁴³⁻⁴⁷. A
696 comparison between experimental and simulated binding modes of GpA is highlighted in the
697 right panel. The experimental structure was taken from solution NMR in micelles (PDB code
698 AFO)⁴⁹. Near identical experimental structures were obtained by ssNMR in nanodiscs and X-
699 ray crystallography in a lipid cubic phase⁴⁹⁻⁵¹.

700
701

702 ONLINE METHODS

703

704 1) CG models

705 CG mappings of small molecules were initially inspired by the standard Martini
706 2 models, when they were available. Due to the well-balanced properties of
707 the regular (R), small (S), and tiny (T) beads in Martini 3, the CG models now
708 follow more specific rules for mapping. For instance, over-representing 3-to-1
709 or 2-to-1 fragments by the usage of R-beads is always avoided. Aromatic
710 rings without substituents are composed of T-beads and, in case of
711 substituents, S-beads are used. Aliphatic rings without substituents are
712 usually based on S-beads, which better reproduce their molecular shape.
713 More technical details about the mapping rules and bead types used are
714 given in the **Supplementary Notes, sections C1 and C3**. As in the previous
715 version of Martini^{5,20,67-69}, bonded parameters are based on atomistic
716 simulations or high-resolution experimental data. The main difference in
717 Martini 3 lies in the protocol to derive bond lengths, which are now based on
718 matching overall volume and shape of the molecules (**Supplementary Notes,**
719 **section C2**). In this spirit, the bonded parameters of the protein models were
720 also slightly modified from the standard Martini 2.2 values^{68,70}, including the
721 addition of side chain corrections⁷¹, based on experimental reference
722 structures. Backbone bead types do not depend on the secondary structure
723 anymore, but are now represented by P2 beads, except for proline (SP1a),
724 alanine (SP2, with an additional bead for the side chain) and glycine (SP1).
725 Adapted versions of Gō-like models⁷² or Elastic Networks⁷³ were used to
726 maintain the tertiary protein structure. All CG protein models were constructed
727 using *Martinize2*, described in **Supplementary Codes, section H1**. Lipid

728 mapping was inspired by the previous Martini model⁷⁴⁻⁷⁷, but now following
729 the Martini 3 rules for mapping and also with adaptations in the bonded
730 parameters inspired by the “extensible model” of Carpenter *et al.*⁷⁸.

731

732 **2) General setup for CG MD simulations and Analysis**

733 Settings for the CG simulations followed, in general, the “new” Martini set of
734 simulation parameters⁷⁹ using the leap-frog algorithm⁸⁰ for integrating the
735 equations of motion. The Verlet neighbor search algorithm⁸¹ is used to update
736 the neighbor list every 20 steps with a buffer tolerance of 0.005 kJ·mol⁻¹·ps⁻¹.
737 For the Lennard-Jones terms, we used a cutoff scheme with a value of 1.1 nm
738 and the Verlet cutoff scheme⁸² for the potential-shift. Long range electrostatic
739 interactions were treated with reaction field⁸³ or Particle Mesh Ewald (PME)⁸⁴,
740 with relative permittivity set to $\epsilon_r=15$ and a cutoff value of 1.1 nm. Reaction
741 field was used for most of the systems, except the ones explicitly mentioning
742 PME. Periodic boundary conditions were used in all three dimensions. For the
743 production simulations, the velocity rescaling thermostat⁸⁵ (coupling time
744 constant of 1.0 ps) and the Parrinello-Rahman barostat⁸⁶ (coupling time
745 constant of 12.0 ps) were employed to maintain temperature and pressure,
746 respectively. Except for equilibration runs, a time step of 20 fs was used for all
747 systems. CG simulation settings are available as input files for GROMACS on
748 the Martini portal <http://cgmartini.nl>. GROMACS 2016.x and 2018.x were
749 used to run all the MD simulations^{87,88}. For automated running and managing
750 the Martini 3 simulations, we provide an updated version of *Martinate*^{89,90},
751 described in **Supplementary Codes, section H2**. All the analysis were
752 performed using *gmx analysis tools* (GROMACS 2016 and 2018)^{87,88}, VMD

753 1.9.4a12⁹¹, xmgrace (5.1.25) and MDAnalysis⁹². The graphs were plotted
754 using Excel 2016, xmgrace (5.1.25) and gnuplot (5.2). Figures were compiled
755 using VMD 1.9.4a12 and Inkscape 1.1.

756

757 **3) Parameter calibration, tests and validation**

758 In order to parametrize the LJ parameters of single beads and also test the
759 Martini 3 CG models, many molecular systems and methods were used in this
760 work. The overall iterative approach was not based in rigorous separation of
761 calibration and validation groups. As Martini is based on pair interactions, it is
762 hard to find simple systems that cover enough points in the interaction matrix
763 for all bead size combinations. So, complex systems are not only used for
764 validation but can be part of the calibration. The tests performed were
765 separated in “tiers”, which represent systems with different level of complexity.
766 In “tier 0”, isolate beads and simple-molecules are mainly used for calibration
767 of LJ parameters, with balance of different bead sizes and thermodynamics
768 data (e.g. liquid-liquid partitioning and miscibility) used as main targets. In the
769 intermediate “tier 1”, bilayer properties are checked, together with qualitative
770 tests, applied to systems like soluble and transmembrane proteins. These
771 qualitative tests are designed as “yes-or-no” questions to evaluate the overall
772 quality of the force field. At the same time, some points in the interaction
773 matrix were also tested and fine-tuned here. In the final “tier 2”, quantitative
774 tests involving complex systems are performed, including comparisons with
775 experimental or atomistic simulation data. Here most of the system are
776 considered validation. For a complete overview of the parametrization
777 strategy used, see the **Supplementary Notes, section B**. The

778 **Supplementary Notes, section D** provide details of the specific systems and
779 methods used in the tests performed to parametrize and validate the new
780 Martini 3 LJ parameters. Further information on research design is available in
781 the Life Sciences Reporting Summary linked to this article.

782

783 **DATA AVAILABILITY**

784 Force-field parameters and procedures (e.g. tutorials) are publicly available at
785 <http://cgmartini.nl>. Simulation Data (e.g. trajectories) supporting the results of
786 this paper are available from the corresponding authors upon reasonable
787 request.

788

789 **CODE AVAILABILITY**

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

793

794

795

796 **REFERENCES**

- 797 67. López, C. A. *et al.* Martini Coarse-Grained Force Field: Extension to
798 Carbohydrates. *J. Chem. Theory Comput.* **5**, 3195–3210 (2009).
- 799 68. Monticelli, L. *et al.* The MARTINI Coarse-Grained Force Field: Extension
800 to Proteins. *J. Chem. Theory Comput.* **4**, 819–834 (2008).
- 801 69. Grunewald, F., Rossi, G., de Vries, A. H., Marrink, S. J. & Monticelli, L.
802 Transferable MARTINI Model of Poly(ethylene Oxide). *J. Phys. Chem. B* **122**,
803 7436–7449 (2018).
- 804 70. de Jong, D. H. *et al.* Improved Parameters for the Martini Coarse-Grained
805 Protein Force Field. *J. Chem. Theory Comput.* **9**, 687–97 (2013).
- 806 71. Herzog, F. A., Braun, L., Schoen, I. & Vogel, V. Improved Side Chain
807 Dynamics in MARTINI Simulations of Protein–Lipid Interfaces. *J. Chem.*
808 *Theory Comput.* **12**, 2446–2458 (2016).
- 809 72. Poma, A. B., Cieplak, M. & Theodorakis, P. E. Combining the MARTINI
810 and Structure-Based Coarse-Grained Approaches for the Molecular Dynamics
811 Studies of Conformational Transitions in Proteins. *J. Chem. Theory Comput.*
812 **13**, 1366–1374 (2017).
- 813 73. Periole, X., Cavalli, M., Marrink, S.-J. & Ceruso, M. A. Combining an
814 Elastic Network With a Coarse-Grained Molecular Force Field: Structure,
815 Dynamics, and Intermolecular Recognition. *J. Chem. Theory Comput.* **5**,
816 2531–2543 (2009).
- 817 74. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries,
818 A. H. The MARTINI force field: coarse grained model for biomolecular
819 simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).

820 75. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. &
821 Marrink, S. J. Computational Lipidomics with *insane*: A Versatile Tool for
822 Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory*
823 *Comput.* **11**, 2144–2155 (2015).

824 76. Melo, M. N., Ingólfsson, H. I. & Marrink, S. J. Parameters for Martini
825 sterols and hopanoids based on a virtual-site description. *J. Chem. Phys.* **143**,
826 243152 (2015).

827 77. López, C. A., Sovova, Z., van Eerden, F. J., de Vries, A. H. & Marrink, S.
828 J. Martini Force Field Parameters for Glycolipids. *J. Chem. Theory Comput.* **9**,
829 1694–1708 (2013).

830 78. Carpenter, T. S. *et al.* Capturing Phase Behavior of Ternary Lipid Mixtures
831 with a Refined Martini Coarse-Grained Force Field. *J. Chem. Theory Comput.*
832 **14**, 6050–6062 (2018).

833 79. de Jong, D. H., Baoukina, S., Ingólfsson, H. I. & Marrink, S. J. Martini
834 straight: Boosting performance using a shorter cutoff and GPUs. *Comput.*
835 *Phys. Commun.* **199**, 1–7 (2016).

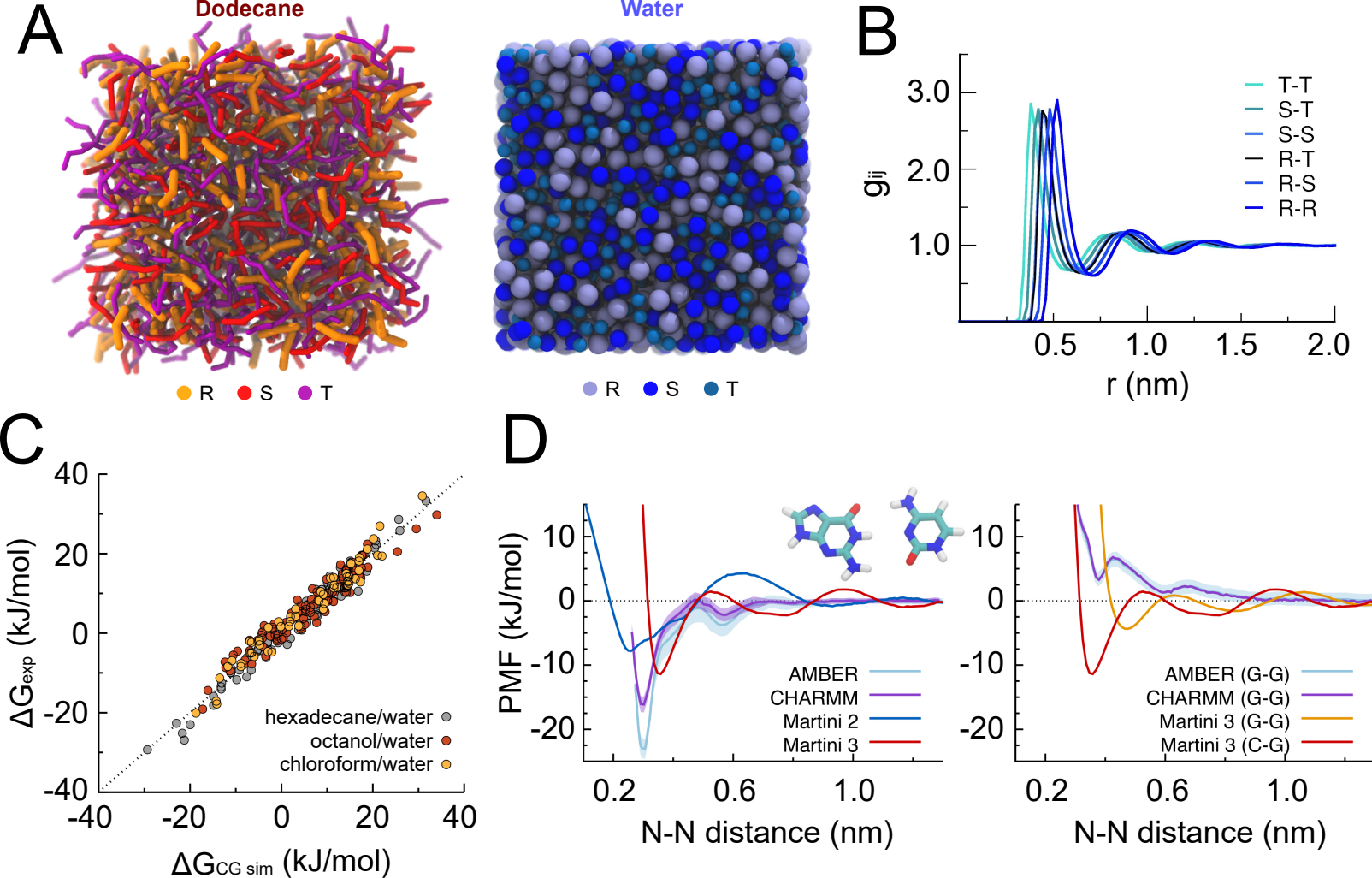
836 80. Hockney, R. W., Goel, S. P. & Eastwood, J. W. Quiet high-resolution
837 computer models of a plasma. *J. Comput. Phys.* **14**, 148–158 (1974).

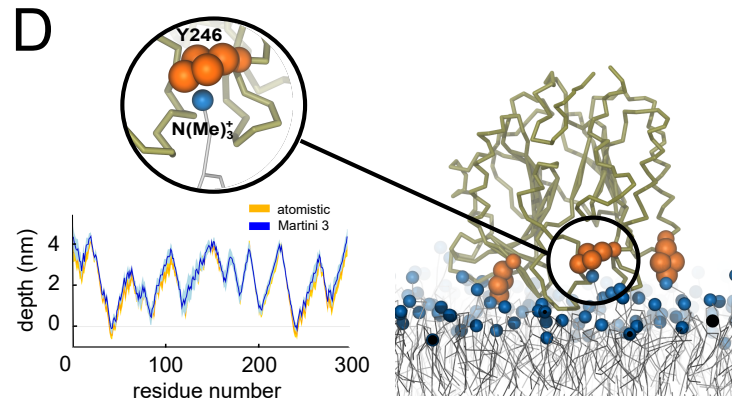
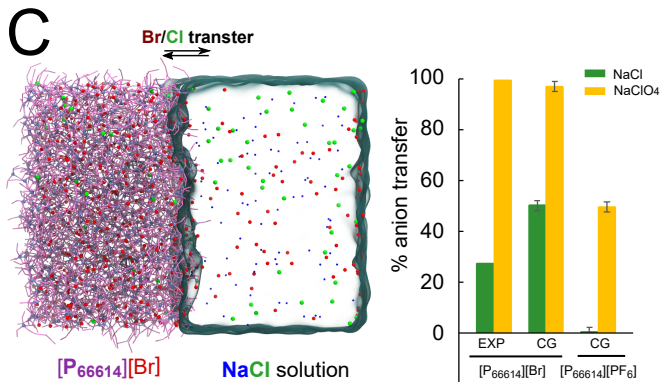
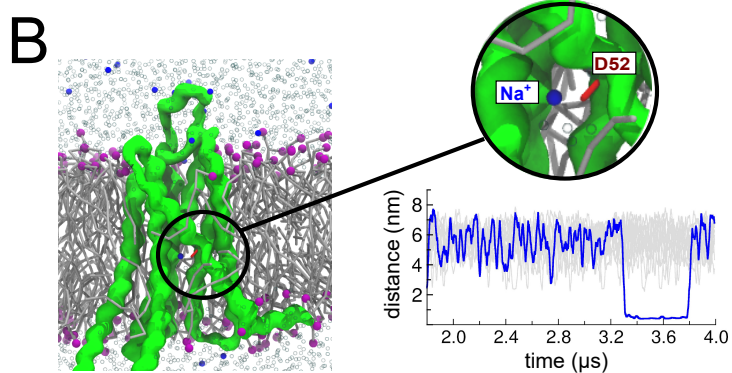
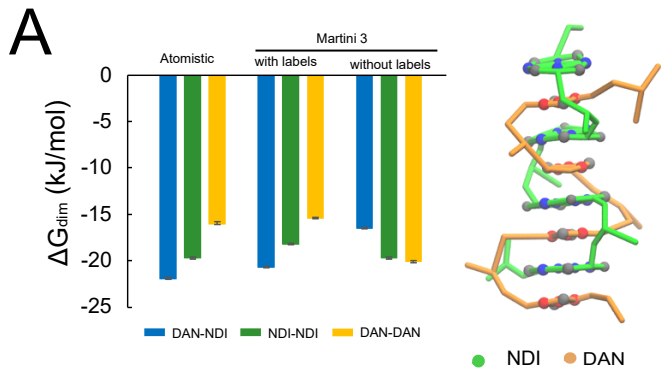
838 81. Páll, S. & Hess, B. A flexible algorithm for calculating pair interactions on
839 SIMD architectures. *Comput. Phys. Commun.* **184**, 2641–2650 (2013).

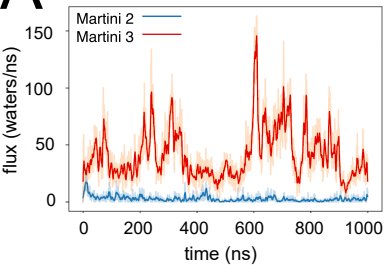
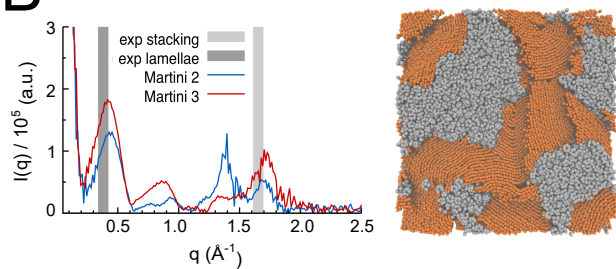
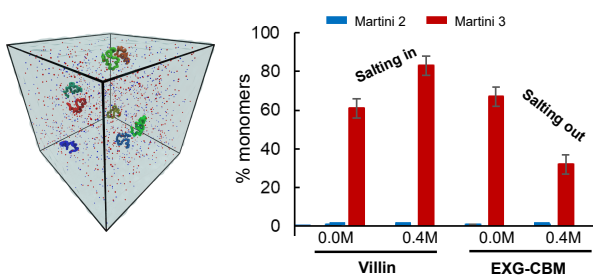
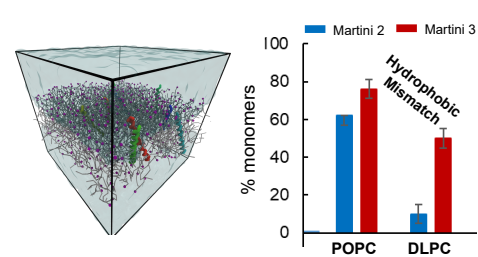
840 82. Verlet, L. Computer ‘experiments’ on classical fluids. I. Thermodynamical
841 properties of Lennard-Jones molecules. *Phys. Rev.* **159**, 98–103 (1967).

842 83. Tironi, I. G., Sperb, R., Smith, P. E. & Van Gunsteren, W. F. A generalized
843 reaction field method for molecular dynamics simulations. *J. Chem. Phys.*
844 **102**, 5451–5459 (1995).

- 845 84. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.*
846 **103**, 8577–8593 (1995).
- 847 85. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through
848 velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
- 849 86. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A
850 new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
- 851 87. Abraham, M. J. *et al.* GROMACS: High performance molecular
852 simulations through multi-level parallelism from laptops to supercomputers.
853 *SoftwareX* **1–2**, 19–25 (2015).
- 854 88. Van Der Spoel, D. *et al.* GROMACS: Fast, flexible, and free. *J. Comput.*
855 *Chem.* **26**, 1701–1718 (2005).
- 856 89. Wassenaar, T. A., Ingólfsson, H. I., Prieß, M., Marrink, S. J. & Schäfer, L.
857 V. Mixing MARTINI: Electrostatic coupling in hybrid atomistic-coarse-grained
858 biomolecular simulations. *J. Phys. Chem. B* **117**, 3516–3530 (2013).
- 859 90. Wassenaar, T. A. *et al.* High-Throughput Simulations of Dimer and Trimer
860 Assembly of Membrane Proteins. The DAFT Approach. *J. Chem. Theory*
861 *Comput.* **11**, 2278–91 (2015).
- 862 91. Humphrey, W., Dalke, A. and Schulten, K. VMD - Visual Molecular
863 Dynamics. *J. Molec. Graph.* **14**, 33–38 (1996).
- 864 92. Gowers, R. J. *et al.* MDAnalysis: A Python Package for the Rapid Analysis
865 of Molecular Dynamics Simulations. In *Proc. 15th Python Sci. Conference* 98–
866 105 (2016).





A**B****C****D****E**