Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Material: Development of Coarse-grained Force Field for Alcohols: an Efficient Meta-Multilinear Interpolation Parameterization Algorithm

Mingwei Wan,[ab] Junjie Song,[b] Ying Yang,[b] Lianghui Gao,*[b] and Weihai Fang[ab]

---

[a] *Institution of Theoretical and Computational Chemistry, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, China*

[b] *Key Laboratory of Theoretical and Computational Photochemistry, Ministry of Education, College of Chemistry, Beijing Normal University, 19 Xin-Jie-Kou-Wai Street, Beijing 100875, China*

*Corresponding author email address: Lianghui Gao (lhgao@bnu.edu.cn)

SI.1. Set of time step for OPT1.35 FF

Table SI-1. Properties of water, butane, ethanol, propanol, butanol, pentanol, hexanol, octanol, and 1,6-hexandiol calculated by using the OPT1.35 FF at time steps from 5 to 50 fs.

	Time step	ρ (kg/m^3)	ΔH_{vap} (kJ/mol)	Γ (mN/m)
W_d-W_d	20 fs	998	45.9	69.0
	30 fs	999	45.9	68.8
	40 fs	999	45.9	69.1
	50 fs	999	45.9	69.4
C_2-C_2	20 fs	583	21.4	12.8
	30 fs	583	21.4	13.2
	40 fs	584	21.4	12.8
	50 fs	584	21.4	13.0
OH-C_2	5 fs	821	42.7	24.0
	20 fs	821	42.8	23.6
	25 fs	822	42.9	23.9
	30 fs	N/A	N/A	N/A
OH-C_3	5 fs	822	46.0	22.0
	20 fs	824	46.1	23.2
	25 fs	824	46.3	23.2
	30 fs	N/A	N/A	N/A
OH-C_2-C_2	5 fs	753	48.5	22.7
	20 fs	754	48.5	22.5
	25 fs	754	48.6	22.0
	30 fs	N/A	N/A	N/A
OH-C_3-C_2	5 fs	792	53.3	24.2
	20 fs	793	53.7	23.9
	25 fs	791	53.1	24.1
	30 fs	790	52.9	N/A
OH-C_3-C_3	5 fs	779	57.3	23.4
	20 fs	779	57.4	24.0
	25 fs	780	57.6	21.7
	30 fs	780	57.6	24.3
OH-C_4-C_4	5 fs	788	66.2	25.9
	20 fs	788	66.1	25.5
	25 fs	788	66.2	25.1
	30 fs	788	66.3	N/A
OH-C_3-C_3-OH	5 fs	1002	95.3	41.3
	20 fs	1001	94.8	40.8
	25 fs	999	94.6	39.5
	30 fs	998	94.3	N/A

To test the efficiency of OPT1.35 FF, simulations in *NPT* and *NVT* ensembles were performed at time steps from 20 to 50 fs, and properties were calculated and compared. As given in Table SI-1, ρ, ΔH_{vap}, and Γ of water (W_d-W_d) and butane (C_2-C_2) calculated at 20 fs to 50 fs are within 4% derivation from the values obtained at 10 fs, indicating that time steps of up to 50 fs does not affect the FF accuracy of water and butane. Similar to OPT FF,[1] the upper limit of time step for propane (C_3) and butane (C_4) is also 50 fs (data not shown). For alcohols, the upper limit of time step is reduced to 25 fs, owing to the small bead size of OH and the short bond length of OH-C. Therefore, a time step of 20 fs is suggested as an efficient and safe step in constant temperature ensembles. As a matter of fact, time steps of 10 to 20 fs are commonly used in many CG methods.[1-4]

We also performed simulations in constant energy (*NVE*) ensemble to evaluate the conservation of energy at different time steps (1, 5, 10, 20, 30, 40, and 50 fs) for water, hexadecane, ethanol, propanol, pentanol, and octanol by using the OPT1.35 FF. The fluctuation, defined as $\Delta E = \langle [E - \langle E \rangle]^2 \rangle^{1/2}$, was calculated for the total energy (E_{tot}) and the potential energy (E_{pot}). First, each of the system was equilibrated in the *NPT* ensemble (298 K and 1 atm.) for 1 ns at a time step of 1 fs. Then, the last frame as well as the velocity distribution from the *NPT* simulation was used as the initial conditions for the *NVE* simulations. For each *NVE* simulation, the total simulation time was 2 ns and the energy information was collected every 0.5 ps. Pair list was updated every 1 step and the pair list cutoff was set equal to the cut off of Morse potential by default. Trajectory in the last 1 ns were used for data analysis. Results are listed in Table SI-2.

Table SI-2. Ratio of average fluctuations in total and potential energy $\Delta E_{\text{tot}}/\Delta E_{\text{pot}}$ for pure water, hexadecane, ethanol, propanol, pentanol, and octanol.

	N_{mol}	**Δ*t* = 1 fs**	**Δ*t* = 5 fs**	**Δ*t* = 10 fs**	**Δ*t* = 20 fs**	**Δ*t* = 30 fs**	**Δ*t* = 40 fs**	**Δ*t* = 50 fs**
W_d-W_d	3200	0.004	0.004	0.013	0.167	0.830	1.863	N/A
C_4-C_4-C_4-C_4	800	0.003	0.002	0.005	0.014	0.030	0.141	1.671
C_4-C_4-C_4-C_4[a]	800	0.001	0.002	0.005	0.018	0.050	0.119	1.649
OH-C_2	1600	0.003	0.101	1.028	N/A	N/A	N/A	N/A
OH-C_3	1600	0.003	0.061	0.934	N/A	N/A	N/A	N/A
OH-C_3-C_2	1100	0.004	0.039	0.627	1.886	N/A	N/A	N/A
OH-C_4-C_4	1100	0.002	0.026	0.544	1.568	N/A	N/A	N/A

[a] *NVE* simulation was performed using a pair list cut off of 1.6 nm.

As discussed in literature, ΔE_{tot} should be less than one fifth of ΔE_{pot} to avoid energy sinks or sources in the *NVE* simulation.[5] To satisfy this relation, the largest time step for water, hexadecane, and alcohols turned to be 20, 40, and 5 fs, respectively (Table SI-2) in our OPT1.35 FF. The polarity from alkane to water to alcohols increased, but CG mapping number decreased, they led to descendent upper limit of time step. It indicates that the safe time step for simulations in *NVE* is shorter than that in NPT ensembles. This is because the absence of friction force in the *NVE* simulation induces a large fluctuation of kinetic energy and total energy. In contrast, the *NPT*/*NVT* simulation conserved energy better by using a heat bath. Therefore, we suggest the usage of a time step of 5 fs in *NVE* ensemble and 20 fs in *NPT*/*NVT* ensemble. We also compared the $\Delta E_{\text{tot}}/\Delta E_{\text{pot}}$ of hexadecane obtained by cutting the pair interaction at 1.60 nm and 1.35 nm. Truncation at 1.60 nm only slightly improved the energy conservation at time steps of 40 and 50 fs. Thus, pair list cutoff of 1.35 nm (equal to the Morse potential cutoff, which is more detailed discussed in the next section) was used for the concern of simulation efficiency.

SI.2 Homotypic W_d-W_d water model and the set of truncation radius

The CSJ 4-to-1 mapped spherical water model[2] used a traditional Morse potential $\varepsilon[e^{\alpha(1-r/R0)}-2e^{0.5\alpha(1-r/R0)}]$ to describe the non-bonded interactions. It utilized a long cut-off R_c of 1.6 nm to reproduce the rational thermodynamic properties. For simplicity, no tail-corrections were introduced for the potential/force. This produced a vapor/liquid surface tension Γ of 71.0 mN/m (Table SI-3), close to the experimental value of 72.0 mN/m.[6] However, when R_c is increased to 2.8 nm, Γ also increases to 89.5 mN/m (Table SI-3). The significant discrepancy indicates a poor continuity of the potential at the truncation radius. The OPT 4-to-1 spherical water model[1] also used a R_c of 1.6 nm and met the same continuity problem.

In contrast, the Martini FF is more efficient via a L-J potential with an R_c of 1.1 nm to 1.2 nm, and applying a shifted/switched function to assure the continuity at the cut-off.[7-8] However, the shifted/switched function varies in different MD codes hindering the accuracy and transferability.[8] In this work, we used a short R_c to ensure efficiency, and did not apply a shifted/switched function to ensure accuracy and transferability. We simply tuned the long-range softness by using a large β to solve the continuity problem. However, our trials show that reducing the cut-off in the OPT 4-to-1 spherical model results in crystallization of water at room temperature similar to the Martini model.[7] This effect implies that the 4-to-1 spherical model might inappropriately mimic a water cluster—especially when the CG potential has a short cut-off radius.

Table SI-3. The continuity of the CSJ and OPT water models tested by tracking vapor/liquid surface tension Γ and $|F_{\mathrm{cut}}|$ with potential truncated at various cut-off radii R_c.

R_c (nm)		**1.2**	**1.3**	**1.6**	**2.0**	**2.5**	**2.8**	**EXP**
Γ (mN/m)	CSJ	31.5	43.8	71.0	85.6	88.9	89.5	72.0
	OPT	24.8	38.3	76.8	107.9	121.6	126.0	
$|F_{\mathrm{cut}}|$ (kJ/mol/nm)	CSJ	1.511	0.882	0.170	0.018	0.002	2×10^{-4}	
	OPT	1.585	1.062	0.307	0.057	0.010	0.002	

Experimental surface tensions (termed EXP) are from ref. 6.

We thus considered a non-spherical and multisite model. In specific, we mapped four water molecules into two bonded particles called a homotypic W_d-W_d model. A harmonic constraint was applied to make the two particles tightly connected. Here we fixed the bond strength k_b (k_b is around 30000 kJ/mol/nm^2 in a constrained bond) and made L_b tunable. We then used five parameters {ε, R_0, α, β, L_b} to reproduce three properties {ρ, ΔH_{vap}, Γ}. We also checked whether the parameter can ensure the continuity of the Morse force at a given cut-off R_c, termed as $|F_{\mathrm{cut}}|$. Only when $|F_{\mathrm{cut}}|$ is lower than a critical value, the potential can be said having a good continuity. For example, in Table SI-3 the Γ from CSJ model at R_c of 2.0 nm is 85.6 mN/m, where $|F_{\mathrm{cut}}|$ is as low as 0.018 kJ/mol/nm. The increasing of R_c up to 2.8 nm results in Γ of 89.5 mN/m, still within 5% of 85.6 mN/m. Therefore, $|F_{\mathrm{cut}}|$ of 0.02 kJ/mol/nm can be set as a criterion assuring potential/force continuity. In this work, our purpose is to reduce R_c and the corresponding $|F_{\mathrm{cut}}|$ as low as 0.02 kJ/mol/nm.

The specific parameterization for water consisted of three parts. In the first step, we used a simplex algorithm[9] to generate a rough initial parameter. In the simplex method, we mainly optimize the merit/error function, the weighted average of deviations of the properties {ρ, ΔH_{vap}, Γ} from experiment:

$$y(P_i) = \frac{1}{3}\left(\left|\frac{\boldsymbol{\rho}_i}{\boldsymbol{\rho}_{exp}} - 1\right| + \left|\frac{\Delta H_{\mathrm{vap}_i}}{\Delta H_{\mathrm{vap}_{exp}}} - 1\right| + |\frac{\Gamma_i}{\Gamma_{exp}} - 1|\right). \qquad \text{(SI-1)}$$

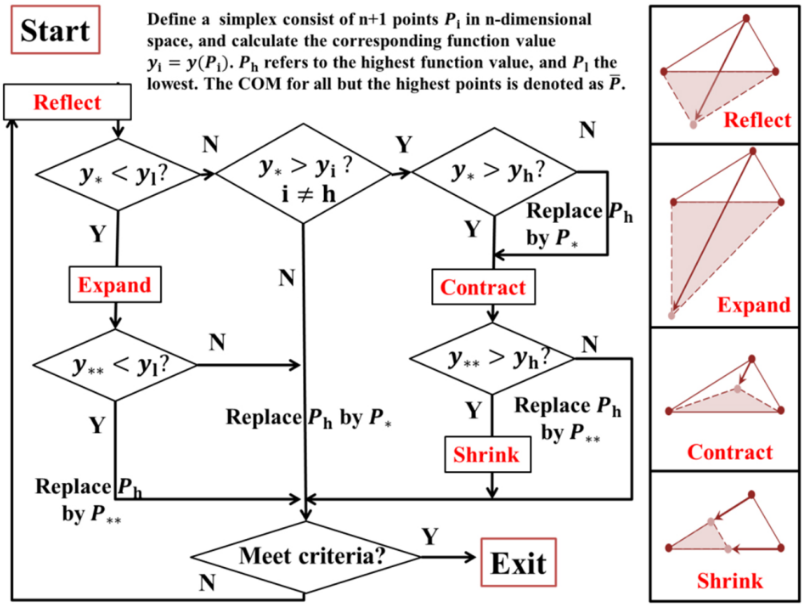

Figure SI-1. The flow diagram of simplex method.

Simplex is a multidimensional optimization algorithm,[9] which is efficient at the early stage of FF parameterization. It has *N*+1 vertices in a *N*-dimensional parameter space. Each vertex corresponds to a parameter set ***P***. To evaluate ***P***, a merit function is defined, for example, as the weighted error of several properties relative to experiment. Simplex has four basic moves: reflect, expand, contract, and shrink. First, the worst point $\boldsymbol{P}_\mathrm{h}$ of all the vertexes is reflected through the centroid $\bar{\boldsymbol{P}} = 1/N \sum_{i=1, i\neq \mathrm{h}}^{N+1} \boldsymbol{P}_i$ of the remaining points, giving a new point $\boldsymbol{P}_* = (1+\gamma_\mathrm{r})\bar{\boldsymbol{P}} - \gamma_\mathrm{r}\boldsymbol{P}_\mathrm{h}$. If $\boldsymbol{P}_*$ is better than all the other vertices, the distance between $\boldsymbol{P}_*$ and $\bar{\boldsymbol{P}}$ is expanded, leading to point $\boldsymbol{P}_{**} = (1-\gamma_\mathrm{e})\bar{\boldsymbol{P}} + \gamma_\mathrm{e}\boldsymbol{P}_*$. If $\boldsymbol{P}_*$ is worse than all the other vertices, the distance between $\boldsymbol{P}_*$ and $\bar{\boldsymbol{P}}$ is contracted, resulting in $\boldsymbol{P}_{**} = (1-\gamma_\mathrm{c})\bar{\boldsymbol{P}} + \gamma_\mathrm{c}\boldsymbol{P}_\mathrm{h}$. If $\boldsymbol{P}_{**}$ is worse than all the other vertices, the simplex is shrunk along the directions from all $\boldsymbol{P}_i$ to $\boldsymbol{P}_\mathrm{l}$: $\boldsymbol{P}_i = 0.5\boldsymbol{P}_i + 0.5\boldsymbol{P}_\mathrm{l}, i \neq \mathrm{l}$. As shown in Figure SI-1, by iterating the 4 moves, the merit function is finally lower than a given threshold value. A detailed usage of simplex algorithm combined with MD can be found in ref. 1.

Table SI-4. Force parameters of water obtained at different stages of optimization.

	P_0	***P_1***	***P_2***	***P_2-T***						
T (K)	298	298	298	243	253	263	273	323	348	373
ε (kJ/mol)	2.321	2.985	3.0145	3.0773	3.0676	3.0574	3.0466	2.9825	2.9431	2.8977
R_0 (nm)	0.479	0.493	0.4969	0.5041	0.5017	0.5000	0.4987	0.4958	0.4953	0.4951
α	7.000	6.224	6.249							
β	7.000	9.289	9.253							
L_b (nm)	0.479	0.493	0.5054							

P_0, *P_1*, *P_2*, and *P_2-T* corresponded to the parameter sets obtained after a quick simplex search, after a Meta-MIP optimization, after a further simplex refinement, and at different temperatures, respectively. The parameters *ε* and R_0 of *P_2-T* are fitted by *ε*(*T*)=0.4949+11.30×exp(-*T*/34.14) and R_0(*T*)=3.245-0.04295×exp(*T*/178.5), respectively.

To set the input parameter of a simplex, we halved ε of the CSJ water parameter. Then, we fixed α and β, let L_b equal to R_0, and only optimize the rest two parameters, ε and R_0, using a 2-dimensional simplex. We used a R_c of 1.2 nm as in many other CG methods.[7-8] After 63 iterations, we obtained a promising parameter set $\boldsymbol{P}_0$ (Table SI-4). The properties predicted by $\boldsymbol{P}_0$ are within 4% deviation from experimental measurement,[10] see Table SI-5. However, the potential has a poor continuity, as reflected by Γ which varied more than 25% when R_c is increased from 1.2 nm to 1.6 nm. The poor continuity is also indicated by the large cut-off force $|F_{1.2}|$ of 0.174 kJ/mol/nm (Table SI-5).

Table SI-5. Continuity of the Morse potentials with various force parameters checked by simulating properties of W_d-W_d, C_3, C_4, and C_2-C_2 via setting different R_c.

R_c (nm)	1.20	1.35	1.60	EXP	1.20	1.35	1.60	EXP
properties		ρ (kg/m^3)				ΔH_{vap} (kJ/mol)		
W_d-W_d P_0	1032		1091	997.0	45.4		47.5	44.0
W_d-W_d P_1	1012	1025	1031	997.0	45.1	45.4	45.5	44.0
W_d-W_d P_2	984	998	1004	997.0	45.5	45.9	46.0	44.0
C_3		494	494	493.1		14.9	14.9	14.8
C_4		572	573	572.9		21.2	21.2	21.0
C_2-C_2		583	583	572.9		21.4	21.4	21.0
properties		Γ (mN/m)				$\|F_{cut}\|$ (kJ/mol/nm)		
W_d-W_d P_0	69.5		86.9	72.0	0.174		0.009	
W_d-W_d P_1	66.3	70.4	71.8	72.0	0.072	0.018	0.002	
W_d-W_d P_2	65.2	69.0	70.5	72.0	0.080	0.020	0.002	
C_3		6.4	6.5	6.6	0.041	0.008	5×10^{-4}	
C_4		11.9	12.1	11.9	0.050	0.008	4×10^{-4}	
C_2-C_2		13.0	13.1	11.9	0.016	0.003	2×10^{-4}	

Experimental properties are from ref. 10.

In the second step, we tried to reduce $|F_{1.2}|$. The simplex is not the best algorithm in this step because the merit function in eqn (SI-1) is inflexible to include the effect of $|F_{1.2}|$.[1, 3] Rather, we used Meta-MIP algorithm wherein $|F_{1.2}|$ and other properties can be handled separately. The $\boldsymbol{P}_0$ obtained in the foregoing step was used as the input parameter set to construct Cell 1 in Meta-MIP. Here $\{\alpha, \beta\}$ was no longer fixed but L_b was still set identical to R_0 because L_b has a lower impact on $|F_{1.2}|$. The dimensions of the $\{\varepsilon, R_0, \alpha, \beta\}$ cell were set to 0.70 kJ/mol, 0.018 nm, 1.000, and 1.000, respectively. 10^4-16=9984 points were interpolated in each cell. The screening condition was set to assure $\{\rho, \Delta H_{vap}, \Gamma\}$ within {5%, 5%, 5%} derivation from experiment,[6] the ranking condition was set to reduce $|F_{1.2}|$, and the extrapolation factor was equal to 0.0.

Figure SI-2 shows that as the Meta-MIP is iterated, the three properties for each $\boldsymbol{P}_b$ in Cells 1-5 stay well within 5% deviation from experiment[6] (dashed lines in Figures SI-2e-g), while $|F_{1.2}|$ decreases rapidly. In Cell 7, the properties of the best parameter set deviate more than 5% from experiment[6] (data not shown), so the Meta-MIP is stopped at Cell 6. However, the $|F_{1.2}|$ of the best parameter set in Cell 6 (called $\boldsymbol{P}_1$, see Table SI-4) is 0.072 kJ/mol/nm, still larger than the threshold value of 0.020 kJ/mol/nm. To balance accuracy and continuity, we extended R_c at the expense of efficiency. The $|F_{cut}|$ of $\boldsymbol{P}_1$ at an R_c of 1.20 nm, 1.25 nm, 1.30 nm, 1.35 nm, 1.40 nm is 0.072, 0.045, 0.028, 0.018, and 0.011 kJ/mol/nm, respectively. Since $|F_{1.35}|$ is right below the threshold value, 1.35 nm is set as the truncation radius. Simulations using $\boldsymbol{P}_1$ at R_c of 1.35 nm give rise to $\{\rho, \Delta H_{vap}, \Gamma\}$ (red dots in Figures SI-2e-g and Table SI-5) satisfying the accuracy demand. Meanwhile, any R_c larger than 1.35 nm is safe for usage without affecting the FF accuracy much (Table SI-5).

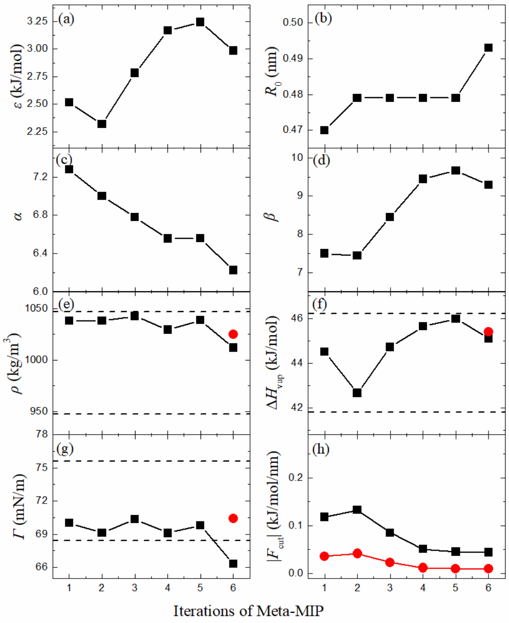

Figure SI-2. Variation of (a) ε, (b) R_0, (c) α, and (d) β of $\boldsymbol{P}_b$ as a function of Meta-MIP iterations, and resulted magnitudes of (e) density, (f) heat of vaporization, (g) gas-liquid surface tension, and (h) Morse force for the W_d-W_d water model. Black squares are obtained by Morse potential truncated at R_c of 1.20 nm, and red dots are obtained with R_c of 1.35 nm. In (e-g), the dashed lines correspond to the 105% and 95% of experimental values.[6]

The final step is to optimize L_b and further refine other parameters using simplex by setting $\boldsymbol{P}_1$ as the initial parameter set. After 49 iterations, we obtained the optimal parameter set $\boldsymbol{P}_2$ (Table SI-4). The predicted values of ρ, ΔH_{vap}, and Γ are 998 kg/m^3, 45.9 kJ/mol, and 69.0 mN/m, respectively (Table SI-5); all deviate less than 4.4% from the experimental values.[6] The Morse force $|F_{1.35}|$= 0.020 kJ/mol/nm indicates that the potential is nearly continuous at a cut-off radius of R_c=1.35 nm (Table SI-5).

To test the temperature stability of force parameter set $\boldsymbol{P}_2$ for the W_d-W_d model, we calculated density ρ, heat of vaporization ΔH_{vap}, surface tension Γ, solvation free energy ΔG_s, self-diffusion coefficient D, isothermal compressibility k_T, and thermal expansion coefficient α_T over a wide temperature range from 243 K to 373 K. Figure SI-3 shows that the properties at different temperatures predicted by $\boldsymbol{P}_2$ agree well with experiments (especially ΔH_{vap}, Γ, and ΔG_s).[6, 11-12] In fact, the performance of $\boldsymbol{P}_2$ is also comparable to the OPT FF.[1]

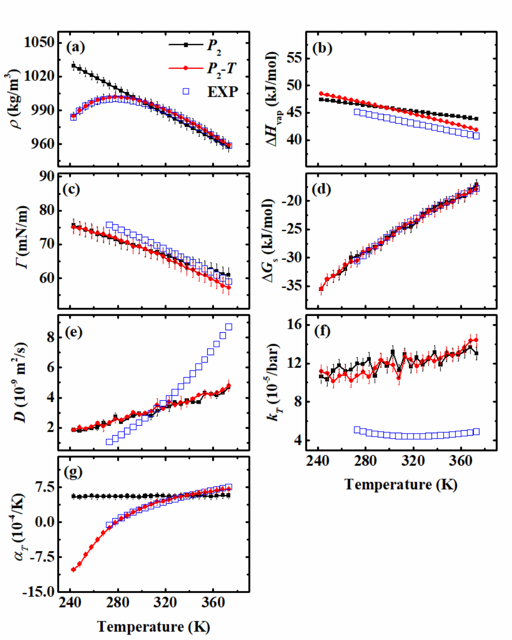

Figure SI-3. Properties of water at various temperatures: (a) density, (b) heat of vaporization, (c) surface tension, (d) solvation free energy, (e) diffusion coefficient, (f) isothermal compressibility, and (g) thermal expansion. The results calculated from ***P***$_2$ (black squares) and ***P***$_{2\text{-}T}$ (red dots) are compared with experimental (EXP) data[6, 11-12] (open blue squares).

A major issue of the CSJ[2] and OPT[1] waters as well as the OPT1.35 model with parameter ***P***$_2$ is that the density anomaly is not captured. The physical water has a maximum ρ at 277 K[6] while the ρ predicted by CSJ/OPT[1-2] and ***P***$_2$ (Figure SI-3a) decrease linearly with temperature. To fix this issue, we used simplex to optimize ε and R_0 at different temperatures while keeping the less sensitive parameters α, β, and L_b the same as those in ***P***$_2$. The optimal parameters at seven additional temperature points are obtained (see Table SI-4). Then, ε/R_0 versus T is least square fitted by an exponential function (given also in Table SI-4). The modified temperature-dependent parameter is termed as ***P***$_2$-T. Figure SI-3a shows that ***P***$_2$-T successfully captures the density anomaly that is maximized at 277 K. In Figure SI-3g, the predicted α_T is almost identical to experimental values.[12] Meanwhile, compared to ***P***$_2$, the accuracy of ΔH_{vap}, Γ, ΔG_{s}, D, and k_T is well preserved.

In a summary, the good performance of both ***P***$_2$ and ***P***$_2$-T parameters indicates that a neutral water model with a complex geometry is promising to use short cut-off while also having good continuity. In contrast, a simple spherical water model fails to simultaneously use short cut-off and have good continuity, unless an additional modification function is used. It should be emphasized that the main reason that we developed the W_d-W_d model is to assure efficiency and continuity. At the moment, it is not convincing to say that W_d-W_d is superior to a spherical model. Nevertheless, there is evidence that W_d-W_d is reasonable. For example, the polarized three-site Martini water[13] has a neutral central site connecting with two charged sites, where the skeleton angle is set as 0 rad. When the two virtual sites overlap, the three-site model becomes two-site. In addition, W_d-W_d is also similar to the titratable Martini water,[14] which is also two-site provided the proton bead is dissociated. Although W_d-W_d is more efficient than a three-site model, the polarized/hydrogen-bonded/pH-dependent version of W_d-W_d should be developed and further evaluated, which is beyond the scope of this work.

SI.3 Alkane model in the OPT1.35 FF

For nonpolar *n*-alkanes, two to four carbon atoms with their attached hydrogen atoms are mapped to one CG C_2, C_3, or C_4 bead. The recently developed OPT FF[1] for C_3 and C_4 nicely reproduce ρ, ΔH_{vap}, and Γ of propane and butane (Table SI-5) with a maximum deviation of 3.1 % from experiment even when the Morse potential is truncated at R_c=1.35 nm. The low $|F_{1.35}|\approx0.008$ kJ/mol/nm shows a good continuity of potential (Table SI-5). Therefore, the OPT parameters[1] (both bonded and non-bonded parameters were thoroughly optimized) for C_3 and C_4 are directly transferred to OPT1.35 FF. Detailed discussions on the properties of *n*-alkanes constructing from C_3 and C_4 are published.[1]

An ethyl group represented by a C_2 bead was not defined and parameterized in both CSJ[2] and OPT[1] FFs. In order to simulate pentane (C_2-C_3), we further parameterize C_2 by simulating butane (C_2-C_2). The bond force constant k_b is set to 1250 kJ/mol/nm^2 consisting with other alkanes,[1] and the bond length is fixed at $L_{b,C_2\text{-}C_2}$= $[2(L_{b,C_3\text{-}C_3})^3-(L_{b,C_4\text{-}C_4})^3]^{1/3}$=0.36 nm based on the approximate relationship of volume, $V_{C_2}=2V_{C_3}-V_{C_4}$. We now have four non-bonded parameters $\{\varepsilon, R_0, \alpha, \beta\}$ to be optimized to match three properties $\{\rho, \Delta H_{vap}, \Gamma\}$. Here, we directly use the Meta-MIP algorithm. The initial parameters of C_2 are approximately extrapolated from C_3 and C_4 via $PA_{C_2\text{-}C_2}=(PA_{C_3\text{-}C_3})^2/PA_{C_4\text{-}C_4}$ where PA stands for ε, α, or β; $R_{0,C_2\text{-}C_2}=[2\times(R_{0,C_3\text{-}C_3})^3-(R_{0,C_4\text{-}C_4})^3]^{1/3}$. Only two iterations can reduce the average deviation to 1.9%. The non-bonded and bonded parameters between C_2 and C_3 or C_4 are obtained via the combination rules $[PA_{C2\text{-}N}=(PA_{C2\text{-}C2}\times PA_{N\text{-}N})^{1/2}$ and $PB_{C_2\text{-}N}=(PB_{C_2\text{-}C_2}+PB_{N\text{-}N})/2$, where PA stands for ε, α, or β; PB stands for R_0 or L_b; N is C_3 or C_4].

To test the accuracy and transferability of the FF for C_2, the thermodynamic properties including ρ, ΔH_{vap}, Γ, ΔG_s, D, k_T, and α_T of butane (C_2-C_2), pentane (C_2-C_3), and hexane (C_2-C_4) are calculated by simulations. Table SI-6 shows that the relative deviations of butane, pentane, and hexane are less than 7.6 %, 3.6 %, and 3.6%, respectively, compared to experimental ρ, ΔH_{vap}, Γ, ΔG_s.[6, 10, 15] For D, k_T, and α_T, which are often poorly predicted by many CG FFs,[16-17] our OPT1.35 FF can give satisfying results on the same order of magnitude of experimental values.[18-20]

Table SI-6. Thermodynamic properties of butane (C_2-C_2), pentane (C_2-C_3), and hexane (C_2-C_4).

alkane	ρ (kg/m^3)	ΔH_{vap} (kJ/mol)	Γ (mN/m)	ΔG_s (kJ/mol)	D (10^{-9} m^2/s)	k_T (10^{-5} bar^{-1})	α_T (10^{-4} K^{-1})
butane [a]	583	21.4	12.8	-12.2	4.7	27.4	16.3
butane [b]	573	21.0	11.9	-11.4	6.9	30.5	19.8
pentane [a]	606	26.4	15.7	-14.7	4.5	21.4	13.0
pentane [b]	621	26.4	15.5	-14.2	5.2	21.8	16.4
hexane [a]	641	31.8	18.6	-17.6	3.5	15.6	10.7
hexane [b]	656	31.6	18.0	-17.0	4.2	16.7	14.1

[a] Properties calculated using the OPT1.35 FF. [b] Experimental data from refs. 6, 10, 15, 18-20.

SI.4 CG OPT1.35 FF between water and alkane

We optimized the FF between water and alkanes by matching the hydration free energies ΔG_{hyd} and interfacial tensions Γ_{int} of alkanes.[21-22] Since only two properties are available, we fixed α and β via the combination rules [$PA_{W_d\text{-}N}=(PA_{W_d\text{-}W_d}\times PA_{N\text{-}N})^{1/2}$, where PA stands for α or β, N is C_2, C_3, or C_4]. Then, we parameterized ε and R_0 for W_d-C_3, W_d-C_4, and W_d-C_2 interactions in sequence. In order to achieve high accuracy, we optimized W_d-C_3 FF by simulating propane/water and nonane/water simultaneously to reproduce {$\Gamma_{\mathrm{int},C_3}$, $\Delta G_{\mathrm{hyd},C_3}$, $\Delta G_{\mathrm{hyd},C_3\text{-}C_3\text{-}C_3}$}. The W_d-C_4 FF is obtained by simulating butane/water and dodecane/water simultaneously to reproduce {$\Gamma_{\mathrm{int},C_4}$, $\Delta G_{\mathrm{hyd},C_4}$, $\Delta G_{\mathrm{hyd},C_4\text{-}C_4\text{-}C_4}$}. The W_d-C_2 FF is obtained by simulating butane/water and hexane/water simultaneously to reproduce {$\Gamma_{\mathrm{int},C_2\text{-}C_2}$, $\Delta G_{\mathrm{hyd},C_2\text{-}C_2}$, $\Delta G_{\mathrm{hyd},C_2\text{-}C_4}$}. The optimal parameters of all are found within two iterations of meta-MIP.

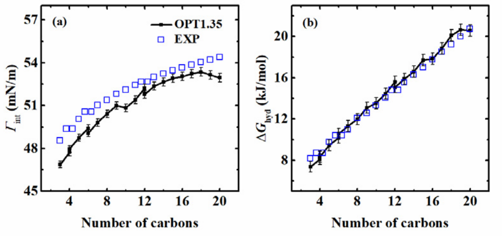

Figure SI-4. Water/alkane interface tension Γ_{int} and hydration free energy ΔG_{hyd} of *n*-alkanes composed of 3 to 20 carbons. Experimental data are from refs. 21-22.

Table SI-7. Interfacial tension Γ_{int} and hydration free energy ΔG_{hyd} of alkanes.

	Γ_{int} (mN/m)		ΔG_{hyd} (kJ/mol)	
Alkanes	OPT1.35	EXP	OPT1.35	EXP
C_3	46.9	48.53	7.4	8.18
C_4	47.7	49.37	8.1	8.70
C_2-C_2	48.0	49.37	8.3	8.70
C_2-C_3	48.7	50.04	9.5	9.76
C_2-C_4	49.4	50.55	10.2	10.40
C_3-C_3	49.1	50.55	10.6	10.40
C_3-C_4	49.8	51.02	11.3	10.96
C_4-C_4	50.4	51.38	12.0	12.10
C_3-C_3-C_3	51.0	51.81	13.1	12.58
C_3-C_3-C_4	50.8	52.11	13.5	13.32
C_3-C_4-C_4	51.4	52.42	14.5	14.06
C_3-C_3-C_3-C_3	52.2	52.67	15.6	14.80
C_4-C_4-C_4	51.8	52.67	15.1	14.80
C_3-C_3-C_3-C_4	52.4	52.99	15.9	15.54
C_3-C_3-C_4-C_4	52.7	53.22	16.6	16.28
C_3-C_3-C_3-C_3-C_3	52.9	53.45	17.7	17.02
C_4-C_4-C_4-C_4	53.0	53.65	17.8	17.76
C_3-C_3-C_3-C_4-C_4	53.2	53.85	18.9	18.50
C_3-C_3-C_3-C_3-C_3-C_3	53.4	54.03	20.1	19.24
C_3-C_3-C_3-C_3-C_3-C_4	53.1	54.21	20.7	19.98
C_3-C_3-C_3-C_3-C_4-C_4	53.0	54.38	20.6	20.72

Data in red indicates a larger than 7 % deviation from experiment.[21-22]

To test the validation of the force parameters, we calculated ΔG_{hyd} and Γ_{int} of *n*-alkanes from propane to eicosane. The mapping details of these alkanes are given in Table SI-7. The results in Figure SI-4 and Table SI-7 show that, except for ΔG_{hyd} of propane (with deviation of 7.4/8.18-1=9.5 %) and butane (with deviation of 8.1/8.70-1=6.9 %), the ΔG_{hyd} and Γ_{int} for *n*-alkanes with lengths of 3 to 20 are within 4.2% of the deviation from experimental values.[21-22] The performance of the OPT1.35 FF is also better than the OPT FF;[1] the latter FF underestimates the ΔG_{hyd} for *n*-alkanes with lengths of 8 to 20. This fact suggests again that the two-site water model might be more suitable than a one-site model even for a water/alkane binary system.

SI.5. Radial distribution functions and structural distributions.

To study the structural details of the CG models in the OPT1.35 FF, we analyzed the radial distribution functions (RDF) of W_d-W_d, C_2-C_2, C_3-C_3, C_4-C_4, OH-OH, OH-C_2, OH-C_3, and OH-C_4. The RDFs in version of OPLS[23] and Martini[7] FFs were also calculated for comparisons. The RDFs from OPT1.35 FF were obtained by using the last 10 ns trajectory of bulk liquids. In AAMD simulations, the geometry structures were obtained from the ATB and Repository Version 3.0,[24-25] and the corresponding OPLS FF[23] was generated using the LigParGen tool.[26-28] 2000 SPC/E water or 500 non-water molecules were initially randomly placed in a cubic box of size of 5×5×5 nm^3, and minimized using the steepest descent method, followed by a 100-ns-long NPT simulation. Temperature was set to 298 K by using the V-rescale thermostat[29] with coupling constant of 0.5 ps, and pressure is set to 1 bar by using the Berendsen barostat[30] with coupling constant of 5.0 ps and compressibility of 4.5×10^{-4} bar^{-1}. Time step was set to 2 fs. Short-range neighbor list, electrostatic, and van der Waals cutoffs were all set to 1.2 nm. For molecules except water, the atomic trajectories were then mapped into CG trajectories. The last 10 ns trajectory consisting 100 frames were used for RDF analysis. For CG Martini water,[7] we simulated 500 W beads (equal to 2000 real water) with the standard NPT (298 K + 1bar) setting.

Figure SI-5a compares the W_d-W_d RDF in OPT1.35 FF, the W-W RDF in Martini FF,[7] and the O-O RDF of SPC/E water.[31] The locations of the peaks in RDFs are not comparable because the mapping schemes are different in these FFs, but the shapes of the RDF curves matter. The Martini RDF shows a significant correlation at long-range, which is never seen in the SPC/E water. Such long-range correlation may result in spontaneously solidification (freezing) at room temperature, as found in many CG models.[3, 7] For two-site W_d-W_d model, the RDF from OPT1.35 FF does not show long-range correlation. As a result, OPT1.35 water does not freeze at room temperature. Moreover, the OPT1.35 RDF is more similar to the SPC/E result than Martini, suggesting a better resemblance with real water.

The C_2-C_2, C_3-C_3, C_4-C_4 RDFs are depicted in Figures SI-5b-d. As expected, the RDF does not depend on the molecular environment. For example, the C_2-C_2 distributions show no distinction in the bulk phase of ethanol (green line in Figure SI-5b) and hexanol (OH-C_2-C_4, grey line in Figure SI-5b). This feature is also noticed for C_3 (Figure SI-5c) and C_4 (Figure SI-5d) using either OPT1.35 (solid line) or OPLS FFs (dash line).[23] This fact demonstrates the portability of OPT1.35 FF. Moreover, the OPT1.35 RDFs agree well with the OPLS results, illustrating its accuracy. It should be noted that the similarity between the OPT1.35 and OPLS RDFs diminishes when the carbon number increases. A possible reason might be the absence of dihedral constrains in CG models.

Figure SI-5e shows that there is a big difference between the OPT1.35 and OPLS[23] OH-OH RDFs. The first peak of the OPT1.35 RDF locates at 0.4 nm, but the OPLS peak appears at 0.3 nm. The closer packing of the hydroxyl groups of alcohols in OPLS FF may attribute to the hydrogen-bonding effect. OPT1.35 FF does not consider hydrogen-bonding and gives rise to looser ethanol clusters. As a consequence, the C-type beads in OPT1.35 FF can permeate into the alcohol clusters and block the gathering of alcohol molecules, or surround a free alcohol. This is reflected by the first RDF peaks of OH-C appearing at shorter distances in Figures SI-5f-h. By contrast, the OPLS OH-C RDFs are fairly flat, indicating a poor attraction between alkyl and hydroxyl groups. Therefore, more realistic CG alcohol model should be considered to account for hydrogen-bonding effects in the future.

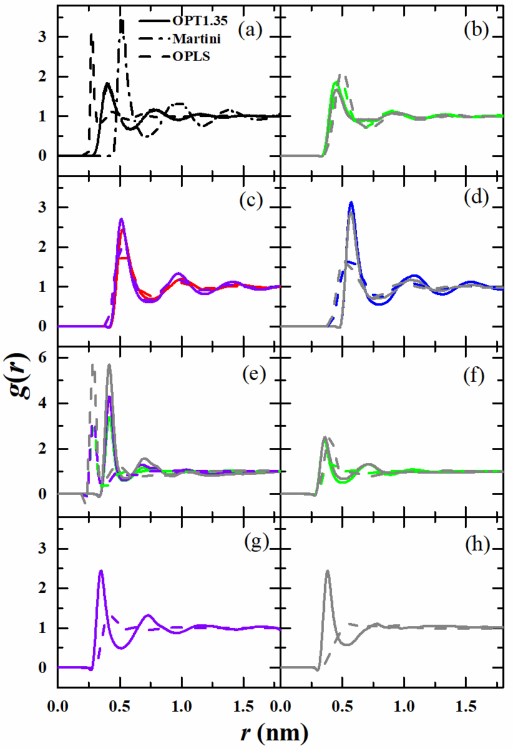

Figure SI-5. The (a) W_d-W_d, (b) C_2-C_2, (c) C_3-C_3, (d) C_4-C_4, (e) OH-OH, (f) OH-C_2, (g) OH-C_3, and (h) OH-C_4 radial distribution functions. The solid, dash dot, and dash lines are results obtained from OPT1.35, Martini,[7] and OPLS[23] FFs, respectively. The black, red, blue, green, purple, and grey colors correspond to the bulk phase simulation of water, propane, butane, ethanol, propanol, and hexanol (OH-C2-C4 in case of the OPT1.35 FF), respectively.

We further calculated the distribution functions of alcohols and compared with that mapped from AA trajectories. The AA MD simulations were performed by using the OPLS FF[23] with the same simulation condition as the RDF analysis. Note that, many FFs[32-33] were bottom-up parameterized (at least partially) by referring to atomistic calculation results. Unfortunately, we found that the harmonic bond length mapped from AA trajectories could not be used herein. As plotted in Figs. SI-6 a-c, the OH-C_2, OH-C_3, and OH-C_4 bond length mapped from AA simulations by using the OPLS FF was around 0.190 nm, 0.212 nm, and 0.266 nm, respectively. By using these values, ethanol, propanol, and butanol had frequent freezing issues. The major reason is that, electrostatic interactions were not explicitly presented, resulting in an excessively ordered packing of alcohols. We did not introduce partial charges in this work, and instead increased the bond length to avoid freezing. For the angle and chain length distributions (Figs. SI-6 d-i), the agreement between CG and AA mapped trajectories increased as the chain length increased, because the effect of OH bead was concealed by the alkane beads. However, large discrepancy still existed. Notice that, even in a bottom-up parameterized CG FF, the CG distribution was often largely different from AA one especially for polar molecules. For example, the CG distribution in the PSO FF[34] was always broad and single peaked, while the AA distribution was narrow and sometimes multiple peaked. Since the current work did not directly train the structural properties, we did not expect the bonded distribution between CG and AA to be perfectly matched. Nevertheless, one could adopt a polarizable model or use a more complicated non-bonded potential to better reproduce both thermodynamic and structural properties.

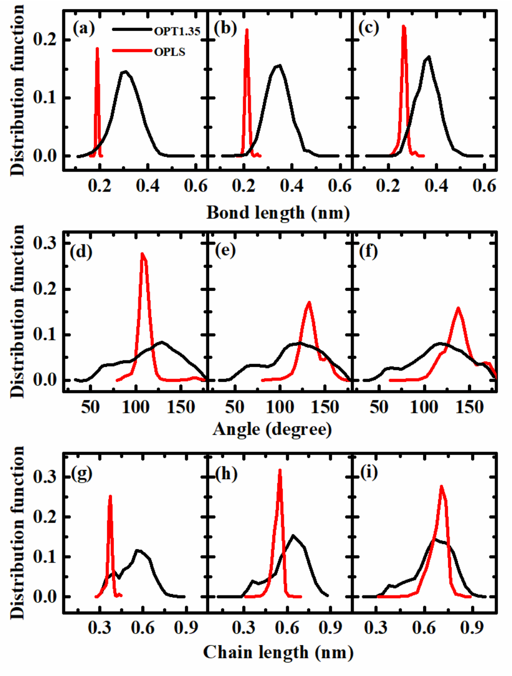

Figure SI-6. The bond distribution function of (a) OH-C_2 in ethanol, (b) OH-C_3 in propanol, and (c) OH-C_4 in octanol; The angle distribution function of (d) OH-C_2-C_2 in butanol, (e) OH-C_3-C_3 in hexanol, and (f) OH-C_4-C_4 in octanol; The chain length distribution function of (g) butanol, (h) hexanol, and (i) octanol. Data in black and red refer to that using the OPT1.35 and OPLS FFs, respectively.

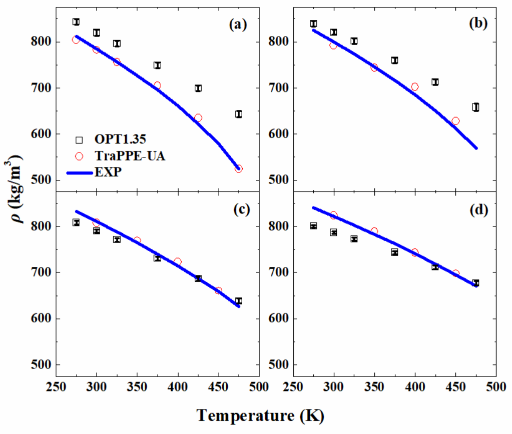

Figure SI-7. Density of (a) ethanol, (b) propanol, (c) pentanol, and (d) octanol at different temperatures. Properties calculated using the OPT1.35 FF, from the TraPPE-UA FF,[35] and from experiment[10] were shown in black squares, red circles, and blue solid lines, respectively.

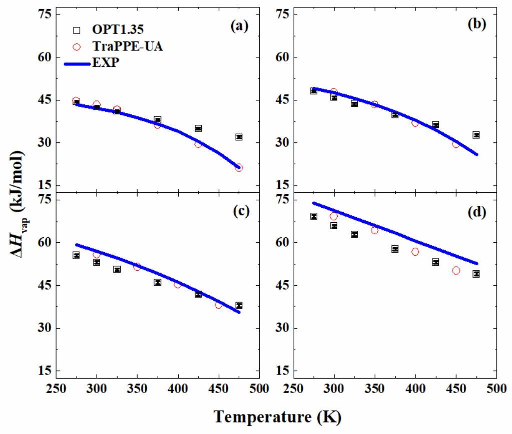

Figure SI-8. Heat of vaporization of (a) ethanol, (b) propanol, (c) pentanol, and (d) octanol at different temperatures. Properties calculated using the OPT1.35 FF, from the TraPPE-UA FF,[35] and from experiment[36] were shown in black squares, red circles, and blue solid lines, respectively.

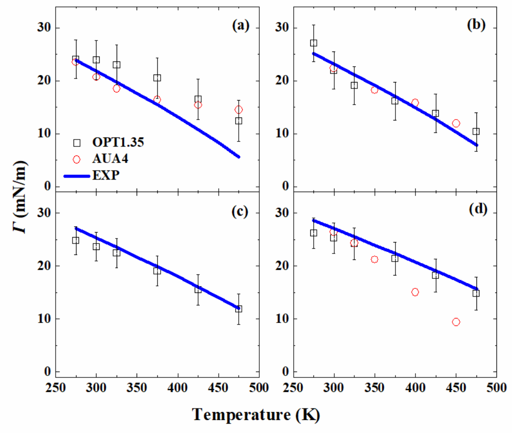

Figure SI-9. Surface tension of (a) ethanol, (b) propanol, (c) pentanol, and (d) octanol at different temperatures. Properties calculated using the OPT1.35 FF, from the TraPPE-UA FF,[35] and from experiment[10] were shown in black squares, red circles, and blue solid lines, respectively.

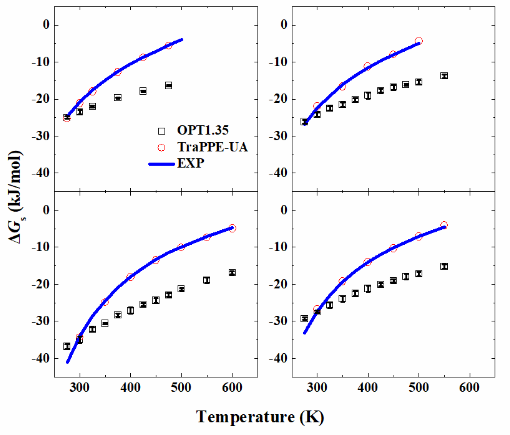

Figure SI-10. Solvation free energy of (a) ethanol, (b) propanol, (c) pentanol, and (d) octanol at different temperatures. Properties calculated using the OPT1.35 FF, from the TraPPE-UA FF,[35] and from experiment[15] were shown in black squares, red circles, and blue solid lines, respectively.

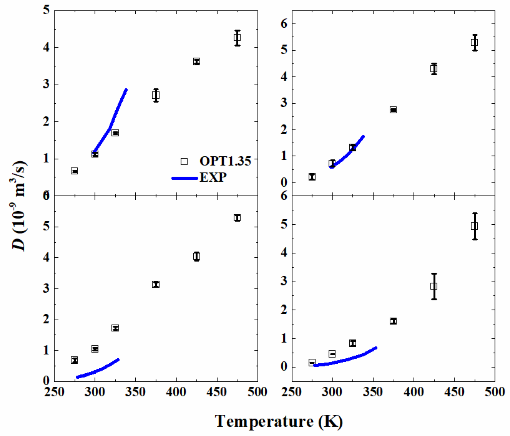

Figure SI-11. Self-diffusion coefficient of (a) ethanol, (b) propanol, (c) pentanol, and (d) octanol at different temperatures. Properties calculated using the OPT1.35 FF and from experiment[37] were shown in black squares and blue solid lines, respectively.

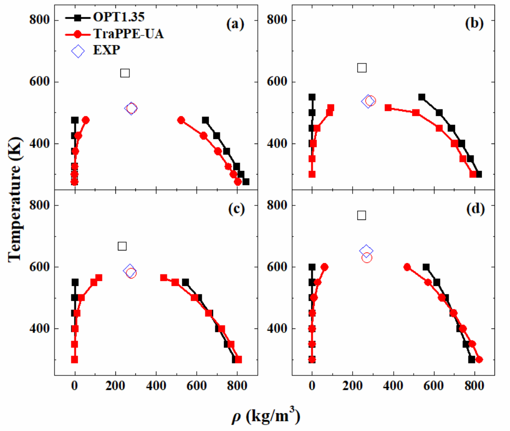

Figure SI-12. The vapour-liquid coexisted curve of (a) ethanol, (b) propanol, (c) pentanol, and (d) octanol. The curve obtained using the OPT1.35 FF and from the TraPPE-UA FF[35] were shown in black and red, respectively. The experimental critical point[38] is given in blue.

Table SI-8. Density ρ, heat of vaporization ΔH_{vap}, surface tension Γ, and solvation free energy ΔG_s of ethanol to dodecanol at 298 K.

	ρ (kg/m^3)		ΔH_{vap} (kJ/mol)		Γ (mN/m)		ΔG_s (kJ/mol)	
Alcohols	OPT1.35	EXP	OPT1.35	EXP	OPT1.35	EXP	OPT1.35	EXP
OH-C_2	821	787.1	42.8	42.32	23.6	22.31	-23.5	-21.03
OH-C_3	824	802.0	46.1	47.45	23.2	23.36	-24.3	-22.81
OH-C_2-C_2	754	805.9	48.5	52.35	22.5	24.16	-25.1	-25.60
OH-C_3-C_2	793	812.0	53.7	57.02	23.9	25.11	-27.3	-27.75
OH-C_3-C_3	779	815.9	57.4	61.61	24.0	25.67	-28.8	-28.95
OH-C_3-C_4	786	819.9	61.7	66.81	26.2	26.44	-31.6	-32.14
OH-C_4-C_4	788	822.9	66.1	70.98	25.5	27.02	-34.6	-34.52
OH-C_3-C_3-C_3	775	824.0	73.4	76.86	29.1	27.97	-36.2	-37.60
OH-C_3-C_3-C_4	781	825.0	78.3	81.50	29.5	28.59	-39.7	-39.58
OH-C_3-C_4-C_4	792	831.0	81.9	85.80	29.5	29.17	-42.0	-43.56
OH-C_4-C_4-C_4	794	829.9	86.6	90.80	29.3	29.75	-45.3	-45.43
OH-C_3-C_3-OH^a	1001	967	94.8	90.20	40.8	42.20	-50.1	-48.40

Data in red indicates a larger than 7 % deviation from experiment.[5-6, 10, 15] The experimental density, heat of enthalpy, solvation free energy of 1,6-hexanediol are from refs.6, 39. The experimental surface tension of 1,6-hexanediol is unavailable, and is extrapolated from the experimental surface tensions of 1,2-ethanediol,[40] 1,3-propanediol,[41] 1,4-butanediol,[6] and 1,5-pentanediol.[42]

Table SI-9. Number of ethanol N_e, water N_w, and pentane N_p in the water-ethanol-pentane system.

c_e	***0.024***	***0.059***	***0.095***	***0.179***	***0.384***	***0.552***	***0.681***
N_e	***192***	***472***	***760***	***1432***	***3072***	***4176***	***5448***
N_w	***1552***	***1482***	***1410***	***1242***	***832***	***556***	***238***
N_p	***1600***	***1600***	***1600***	***1600***	***1600***	***1600***	***1600***

Note that the number of real water molecule is 4N_w, thus the mole fraction of ethanol c_e is $N_e/(N_e+4N_w+N_p)$.

References

1. M. Wan, J. Song, W. Li, L. Gao, W. Fang, Development of Coarse-Grained Force Field by Combining Multilinear Interpolation Technique and Simplex Algorithm. *J. Comput. Chem.* **2020,** *41*, 814-829.
2. S. W. Chiu, H. L. Scott, E. Jakobsson, A Coarse-Grained Model Based on Morse Potential for Water and *n*-Alkanes. *J. Chem. Theory Comput.* **2009,** *6* (3), 851-863.
3. J. C. Fogarty, S. W. Chiu, P. Kirby, Automated Optimization of Water-Water Interaction Parameters for a Coarse-Grained Model. *J. Phys. Chem. B* **2014,** *118* (6), 1603−1611.
4. K. K. Bejagam, S. Singh, Y. An, C. Berry, PSO-Assisted Development of New Transferable Coarse-Grained Water Models. *J. Phys. Chem. B* **2018,** *122* (6), 1958−1971.
5. C. Wohlfarth, B. Wohlfarth, *Pure Liquids: Data Surface Tension of Pure Liquids and Binary Liquid Mixtures*. Springer: Berlin Heidelberg, 2008.
6. J. Rumble, *CRC Handbook of Chemistry and Physics*. 97 ed.; CRC press: Boca Raton, 2017.
7. S. J. Marrink, H. J. Risselada, S. Yefimov, The Martini Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007,** *111* (27), 7812-7824.
8. D. H. de Jong, S. Baoukina, H. I. Ingólfsson, S. J. Marrink, Martini straight: Boosting performance using a shorter cutoff and GPUs. *Comput. Phys. Commun.* **2016,** *199*, 1-7.
9. Vetterling, W. T., *Numerical recipes in Fortran the art of scientific computing*. Cambridge University Press: Cambridge, 1992.

10. C. L. Yaws, *Thermophysical properties of chemicals and hydrocarbons*. 2 ed.; William Andrew: Beaumont, 2014.

11. M. Holz, S. Heil, A. Sacco, Temperature-dependent Self-diffusion Coefficients of Water and Six Selected Molecular Liquids for Calibration in Accurate 1H NMR PFG Measurements. *Phys. Chem. Chem. Phys.* **2000,** *2* (20), 4740-4742.

12. A. Hvidt, K. Rasmussen, T. E. Eriksen, Density, Thermal Expansivity, and Compressibility of Liquid Water from 0" to 150°C: Correlations and Tables for Atmospheric Pressure and Saturation Reviewed and Expressed on 1968 Temperature Scale. *J. Chem. Eng. Data* **1975,** *20* (1), 97-105.

13. S. O. Yesylevskyy, L. V. Schafer, D. Sengupta, S. J. Marrink, Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *PLoS Comput. Biol.* **2010,** *6*, e1000810.

14. F. Grünewald, P. C. T. Souza, H. Abdizadeh, J. Barnoud, A. H. de Vries, S. J. Marrink, Titratable Martini model for constant pH simulations. *J. Chem. Phys.* **2020,** *153*, 024118.

15. C. L. Yaws, *The Yaws handbook of vapor pressure: Antoine coefficients*. 2 ed.; Gulf Professional Publishing: Houston, 2015.

16. A. P. Eichenberger, W. Huang, S. Riniker, Supra-Atomic Coarse-Grained GROMOS Force Field for Aliphatic Hydrocarbons in the Liquid Phase. *J. Chem. Theory Comput.* **2015,** *11* (7), 2925-2937.

17. Y. An, K. K. Bejagam, S. A. Deshmukh, Development of New Transferable Coarse-Grained Models of Hydrocarbons. *J. Phys. Chem. B* **2018,** *122* (28), 7143-7153.

18. D. W. McCall, D. C. Douglass, Diffusion in Liquids. *J. Chem. Phys.* **1959,** *31* (6), 1555-1557.

19. D. Hopfe, *Data Compilation of FIZ. CHEMIE*. Germany, 1990.

20. G. Liessman, W. Schmidt, S. Reiffarth, *Data Compilation of the Saechsische Olefinwerke Boehlen*. Germany, 1995.

21. S. Cabani, P. Gianni, V. Mollica, L. Lepori Group Contributions to the Thermodynamic Properties of Non-Ionic Organic Solutes in Dilute Aqueous Solution. *J. Solution Chem.* **1981,** *10* (8), 563-595.

22. S. Zeppieri, J. Rodríguez, Interfacial Tension of Alkane + Water Systems. *J. Chem. Eng. Data* **2001,** *46* (5), 1086-1088.

23. M. J. Robertson, J. Tirado-Rives, W. L. Jorgensen, Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* **2015,** *11* (7), 3499-3509.

24. A. K. Malde, An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J. Chem. Theory Comput.* **2011,** *7* (12), 4026-4037.

25. M. Stroet, Automated Topology Builder Version 3.0: Prediction of Solvation Free Enthalpies in Water and Hexane. *J. Chem. Theory Comput.* **2018,** *14* (11), 5834-5845.

26. W. L. Jorgensen, J. Tirado-Rives, Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *12*, 6665-6670.

27. L. S. Dodda, J. Z. Vilseck, J. Tirado-Rives, W. L. Jorgensen, 1.14*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations. *J. Phys. Chem. B* **2017,** *121* (15), 3864-3870.

28. L. S. Dodda, I. Cabeza de Vaca, J. Tirado-Rives, W. L. Jorgensen, LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* **2017,** *45*, W331-W336.

29. G. Bussi, Canonical sampling through velocity rescaling. *J Chem Phys.* **2007,** *126* (1), 014101.

30. H. J. C. Berendsen, J. P. M. Postma, Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.

31. H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987,** *91*, 6269−6271.

32. Y. An, K. K. Bejagam, S. A. Deshmukh, Development of Transferable Nonbonded Interactions between Coarse-Grained Hydrocarbon and Water Models. *J. Phys. Chem. B* **2019,** *123* (4), 909−921.

33. W. Shinoda, R. DeVane, M. L. Klein, Multi-property Fitting and Parameterization of a Coarse Grained Model for Aqueous Surfactants. *Mol. Simulat.* **2007,** *33* (1-2), 27-36.

34. O. Conway, Y. An, K. K. Bejagam, S. A. Deshmukh, Development of transferable coarse-grained models of amino acids. *Mol. Syst. Des. Eng.* **2020,** *5*, 675.

35. B. Chen, Jeffrey J. Potoff, J. Ilja Siepmann, Monte Carlo Calculations for Alcohols and Their Mixtures with Alkanes. Transferable Potentials for Phase Equilibria. 5. United Atom Description of Primary, Secondary, and Tertiary Alcohols. *J. Phys. Chem. B* **2001,** *105*, 3093-3104.

36. J. S. Chickos, W. E. Acree, Enthalpies of Vaporization of Organic and Organometallic Compounds, 1880–2002. *J. Phys. Chem. Ref. Data* **2002,** *32*, 519.

37. J. Winkelmann, *Diffusion in gases, liquids and electrolytes Nonelectrolyte Liquids and Liquid Mixtures - Part 2: Liquid Mixtures*. Springer: New York, 2018.

38. B. D. Smith, R. Srivastava, *Thermodynamic Data for Pure Compounds: Part B Halogenated Hydrocarbons and Alcohols*. Elsevier: Amsterdam, 1986.

39. National Center for Biotechnology Information (2020). PubChem Compound Summary for CID 12374, Hexane-1,6-diol. . https://pubchem.ncbi.nlm.nih.gov/compound/Hexane-1_6-diol.

40. National Center for Biotechnology Information (2020). PubChem Compound Summary for CID 174, 1,2-Ethanediol. https://pubchem.ncbi.nlm.nih.gov/compound/1_2-Ethanediol.

41. A. D. John, *Lange's handbook of chemistry*. McGraw-Hill: New York, 1998.

42. surface tension values of some common test liquids for surface energy analysis. http://www.surface-tension.de/.