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

Development of Accurate Coarse-Grained Force Fields for Weakly Polar Groups by an Indirect Parameterization Strategy

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

S1. Direct Parameterization for Molecules from B01 to B04

The bulk phase simulation of a nonpolar molecule (B01 to B04, as listed in Table 1 in the main body) modeled as single CG bead did not meet freezing problem at room temperature. Therefore, the B~B, B~C_4, B~C_3, and B~W_d interactions could be parameterized individually. Here we took *cis*-2-butene (termed as B03) as an example to introduce the optimization procedure for the B~B and B~C_4 interaction parameters. The B03~B03 interaction parameters {ε, R_0, α, β}$_{B03\sim B03}$ were optimized by using the meta-MIP algorithm to reproducing the ρ, ΔH_{VAP}, and Γ of B03. In the first local database Cell 1, sixteen (2^4) neighboring parameter sets were chosen to encircle the initial parameter set $\boldsymbol{P}_S$={4.59 kJ/mol, 0.56 nm, 9, 15} (see Table S4). The dimension of the cell is set as {$\varepsilon_{B03\sim B03}$×2%, $R_{0,\,B03\sim B03}$×2%, $\alpha_{B03\sim B03}$×5%, $\beta_{B03\sim B03}$×5%}. Following the simulation evaluations of these vertexes, the cell resolution was raised by interpolating 10^4-2^4=9984 points. The normal setting of SC={200%, 200%, 200%} combined with RC={lowering the average deviation} was used to concurrently optimize all of the three properties. In four iterations of meta-MIP, an acceptable set of parameters were obtained, which gave a mean error of less than 2% (Figure S1) compared to experimental data.[1-4] In the next step, {ε, R_0, L_0}$_{B03\sim C_4}$ were optimized by matching the ρ, ΔH_{VAP}, and Γ of B03-C_4 molecule as well as the free energy $\Delta G_{C_{16}}$ (if experimental available) of B03 solvated in hexadecane. In this step, {α, β}$_{B03\sim C_4}$ were obtained by using the combination rule Eq. (5) in the article. The rest B03~C_3 and B03~W_d interactions were parameterized in the same way as weakly polar molecules B05 to B16. The optimal CG FFs for B01 to B04 were also given in Table 2.

[a] *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*
[b] *Institution of Theoretical and Computational Chemistry, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, China*
*Corresponding author email address: Lianghui Gao (lhgao@bnu.edu.cn)

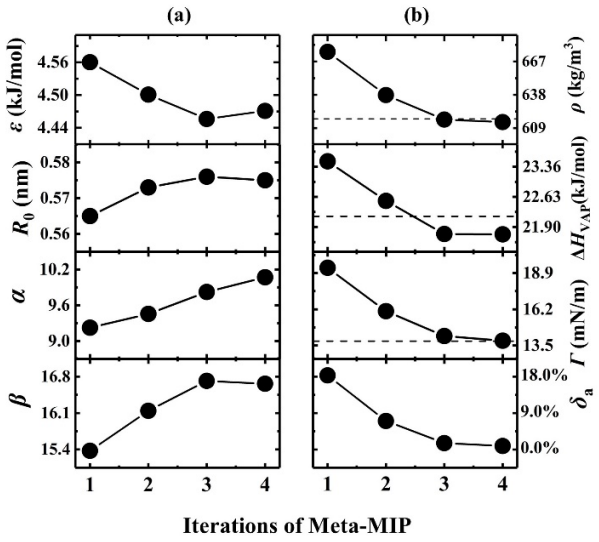

Figure S1 (a) Variations of the force parameters $\varepsilon_{B\sim B}$, $R_{0,B\sim B}$, $\alpha_{B\sim B}$, and $\beta_{B\sim B}$ as a function of the iteration of Meta-MIP for *cis*-2-butene. (b) Corresponding properties ρ, ΔH_{VAP}, and Γ, as well as the mean error δ_a predicted by the force parameters as Meta-MIP proceeds. Experimental data are given by dotted lines.[1-4]

Table S1 Morse force $|F_{1.35}|$ (kJ/mol/nm) at truncation radius of 1.35 nm.

CG-Name	$\|F_{1.35}\|$(B~B)	$\|F_{1.35}\|$(B~C_4)	$\|F_{1.35}\|$(B~C_3)	$\|F_{1.35}\|$(B~W_d)
B01	2.760E-03	2.100E-03	2.270E-03	8.150E-03
B02	2.600E-03	2.360E-03	2.300E-03	7.540E-03
B03	2.830E-03	3.220E-03	2.650E-03	7.840E-03
B04	4.790E-03	3.200E-03	3.340E-03	1.181E-02
B05	3.440E-03	4.230E-03	2.960E-03	8.510E-03
B06	7.338E-05	1.110E-03	2.460E-03	2.040E-03
B07	1.024E-02	8.420E-03	8.970E-03	1.294E-02
B08	1.190E-03	2.760E-03	1.990E-03	6.470E-03
B09	9.210E-03	7.980E-03	8.320E-03	1.251E-02
B10	3.871E-04	1.740E-03	2.000E-03	4.300E-03
B11	7.922E-05	1.090E-03	9.985E-04	2.750E-03
B12	2.100E-03	6.560E-03	7.320E-03	7.430E-03
B13	9.320E-04	2.681E-03	3.671E-03	5.605E-03
B14	3.270E-03	4.810E-03	7.460E-03	8.510E-03
B15	3.180E-03	3.130E-03	3.170E-03	9.430E-03
B16	6.419E-04	1.900E-03	1.350E-03	4.400E-03

Table S2 Thermodynamic properties calculated at various time steps.

CG	ρ (kg/m^3)				ΔH_{VAP} (kJ/mol)				Γ (mN/m)			
Name	10fs	30fs	40fs	50fs	10fs	30fs	40fs	50fs	10fs	30fs	40fs	50fs
B02-C_4	720	719	719	718	39.2	39.2	39.2	39.2	20.1	20.8	20.7	20.8
B02-C_3	712	712	711	710	35.4	35.3	35.3	35.3	19.8	19.8	20.3	19.8
B03-C_4	719	718	718	717	39.9	39.9	39.9	39.8	20.6	21.0	21.3	21.0
B03-C_3	719	719	718	718	35.6	35.6	35.6	35.5	20.0	20.1	20.1	20.1
B04-C_4	769	769	768	768	40.3	40.3	40.3	40.2	22.1	22.5	21.9	22.5
B04-C_3	761	761	760	760	36.1	36.1	36.0	36.0	21.8	21.4	21.5	21.4
B05-C_4	749	749	749	748	43.3	43.2	43.2	43.2	25.1	25.5	25.8	25.5
B05-C_3	763	762	762	761	38.3	38.2	38.2	38.2	24.5	24.4	24.5	24.4
B06-C_4	879	879	878	878	42.4	42.3	42.4	42.3	24.6	24.4	24.5	24.4
B06-C_3	864	863	862	862	38.5	38.4	38.5	38.4	23.6	23.8	23.4	23.8
B07-C_4	862	862	862	861	47.0	47.0	46.9	46.9	27.4	27.7	27.7	27.7
B07-C_3	865	865	865	864	42.2	42.2	42.1	42.1	27.0	26.8	26.6	26.8
B08-C_4	801	800	800	800	47.1	47.0	47.0	46.9	27.0	26.9	26.3	26.9
B08-C_3	823	823	822	821	42.1	42.0	42.0	42.0	26.7	26.6	26.5	26.6
B09-C_4	1124	1123	1123	1123	48.2	48.1	48.1	48.1	27.9	27.8	27.8	27.8
B09-C_3	1165	1165	1165	1164	43.8	43.8	43.8	43.8	27.8	27.8	27.6	27.8
B10-C_4	810	810	809	809	46.8	46.8	46.8	46.7	25.1	25.5	24.8	25.5
B10-C_3	812	812	811	810	42.9	42.9	42.8	42.8	25.4	25.2	25.1	25.2
B11-C_4	761	761	760	760	49.5	49.4	49.4	49.3	25.6	24.3	24.8	24.3
B11-C_3	763	762	762	761	44.7	44.6	44.6	44.5	24.5	24.7	24.7	24.7
B12-C_4	1170	1170	1169	1169	48.4	48.3	48.3	48.3	28.3	27.9	28.3	27.9
B12-C_3	1217	1216	1216	1215	43.8	43.8	43.8	43.8	27.3	28.0	27.5	28.0
B13-C_4	836	836	836	835	48.5	48.5	48.5	48.4	26.4	26.3	26.4	25.5
B13-C_3	828	828	828	827	44.4	44.4	44.4	44.3	25.7	26.1	25.1	25.7
B14-C_4	849	848	848	848	49.7	49.6	49.6	49.6	27.5	27.9	28.0	27.9
B14-C_3	831	831	831	830	45.2	45.2	45.2	45.2	26.6	27.2	26.8	27.2
B15-C_4	795	795	794	794	51.2	51.2	51.1	51.1	28.8	28.7	28.4	28.7
B15-C_3	795	795	794	793	46.6	46.6	46.5	46.5	28.6	27.9	28.3	27.9
B16-C_4	928	927	927	926	49.5	49.5	49.5	49.4	27.8	27.6	27.8	27.6
B16-C_3	985	984	984	983	44.9	44.8	44.8	44.8	28.6	28.4	28.1	28.4

Data in black, blue, and green mean the relative deviation from experiment (ref [1-4]) is in the range of $\delta < 2.5\%$, $2.5\% \leq \delta < 5.0\%$, and $5.0\% \leq \delta \leq 7.0\%$, respectively.

Table S3 Effect of time step on the ratio of the average fluctuations in total and potential energy $\Delta E_T/\Delta E_P$ for pure B01, B02, B03, and B04 bulks. Results were obtained from simulations in NVE ensemble containing 3200 particles. Simulations were performed for 2 ns simulations and the last 1 ns was used for analyses.

		B01, NVE			B02 NVE	
Time steps	ΔE_P	ΔE_T	$\Delta E_T/\Delta E_P$	ΔE_P	ΔE_T	$\Delta E_T/\Delta E_P$
1fs	120.2	0.3	0.002	121.1	0.3	0.003
5fs	118.6	0.5	0.004	121.0	0.4	0.003
10fs	121.0	2.8	0.023	117.5	2.1	0.018
20fs	121.3	14.8	0.122	119.8	17.9	0.149
30fs	126.6	63.8	0.504	139.8	132.3	0.946
40fs	182.7	258.8	1.417	237.3	395.8	1.668
50fs	477.9	934.8	1.956	628.3	1226.7	1.952
		B03, NVE			B04, NVE	
Time steps	ΔE_P	ΔE_T	$\Delta E_T/\Delta E_P$	ΔE_P	ΔE_T	$\Delta E_T/\Delta E_P$
1fs	122.2	0.4	0.003	122.2	0.5	0.004
5fs	123.6	0.3	0.002	119.9	0.4	0.003
10fs	121.4	1.3	0.011	122.8	0.7	0.006
20fs	122.4	15.3	0.125	124.5	7.0	0.056
30fs	131.6	87.9	0.668	130.1	55.9	0.430
40fs	180.8	244.3	1.352	143.9	132.8	0.922
50fs	465.9	902.3	1.937	238.6	390.5	1.637

Table S4 Initial force parameters for sixteen studied molecules or functional groups.

Type	$L_{0,\text{B-C}_4}$ (nm)	$\varepsilon_{\text{B~B}}$ (kJ/mol)	$R_{0,\text{B~B}}$ (nm)	$\alpha_{\text{B~B}}$	$\beta_{\text{B~B}}$	Type	$L_{0,\text{B-C}_4}$ (nm)	$\varepsilon_{\text{B~B}}$ (kJ/mol)	$R_{0,\text{B~B}}$ (nm)	$\alpha_{\text{B~B}}$	$\beta_{\text{B~B}}$
B01	0.44	4.32	0.55	9	15	B09	0.44	6.25	0.56	9	15
B02	0.44	4.44	0.56	9	15	B10	0.43	6.42	0.52	9	15
B03	0.44	4.59	0.56	9	15	B11	0.44	6.48	0.54	9	15
B04	0.44	4.73	0.55	9	15	B12	0.43	6.48	0.54	9	15
B05	0.43	5.53	0.53	9	15	B13	0.44	6.60	0.56	9	15
B06	0.42	5.88	0.49	9	15	B14	0.44	6.61	0.56	9	15
B07	0.44	5.88	0.56	9	15	B15	0.42	7.47	0.51	9	15
B08	0.43	6.14	0.52	9	15	B16	0.41	7.93	0.47	9	15

The initial bond lengths $L_{0,\text{B-C}_4}$ were estimated by using the relations of $L_{0,\text{B-B}}=L_{0,\text{C}_4\text{-C}_4}\times R_{0,\text{B~B}}/R_{0,\text{C}_4\text{~C}_4}$ and $L_{0,\text{B-C}_4}=(L_{0,\text{B-B}}+L_{0,\text{C}_4\text{-C}_4})/2$. Since the B-B dimer might not stably exist, the bond length $L_{0,\text{B-B}}$ only served as an intermediate parameter herein for the convenience of calculation.

Table S5 Thermodynamic properties predicted by a large system performed for 1000 ns.

Molecules	ρ (kg/m^3)			ΔH_{VAP} (kJ/mol)			Γ (mN/m)		
	Sim	Exp	δ (%)	Sim	Exp	δ (%)	Sim	Exp	δ (%)
B02-C_4	719	716	0.4	39.2	40.2[b]	2.5	20.4	20.9	2.4
B02-C_3	712	697	2.2	35.4	36.3[a]	2.5	19.8	20.5	3.4
B03-C_4	718	720	0.3	39.9	40.2[b]	0.7	20.9	22.1	5.4
B03-C_3	719	703	2.3	35.6	36.3[a]	1.9	20.2	20.9	3.3
B04-C_4	769	768[*]	0.1	40.3	41.5[a*]	2.9	22.2	22.2[*]	0.0
B04-C_3	760	750[*]	1.3	36.1	36.6[a*]	1.4	21.4	20.8[*]	2.9
B05-C_4	749	755	0.8	43.3	43.2	0.2	25.3	26.7	5.2
B05-C_3	762	745	2.3	38.3	38.9	1.5	24.7	26.0	5.0
B06-C_4	879	883[*]	0.5	42.4	44.1[a*]	3.9	24.6	25.7[*]	4.3
B06-C_3	863	890[*]	3.0	38.5	40.3[a*]	4.5	23.7	24.6[*]	3.7
B07-C_4	862	872	1.1	47.0	47.7[a]	1.5	27.5	27.0	1.9
B07-C_3	865	874	1.0	42.2	42.8[a]	1.4	26.8	26.2	2.3
B08-C_4	801	813	1.5	47.0	48.7[b]	3.5	27.0	26.3	2.7
B08-C_3	823	810	1.6	42.0	42.3[b]	0.7	26.5	27.9	5.0
B09-C_4	1124	1122	0.2	48.1	47.4	1.5	27.8	29.0	4.1
B09-C_3	1165	1160	0.4	43.8	44.3	1.1	27.8	28.7	3.1
B10-C_4	810	810	0.0	46.8	47.2[a]	0.8	25.3	25.8	1.9
B10-C_3	812	807	0.6	42.9	43.1[a]	0.5	25.2	25.3	0.4
B11-C_4	761	772	1.4	49.5	50.0[a]	1.0	25.3	24.0	5.4
B11-C_3	763	761	0.3	44.7	45.1[a]	0.9	24.5	25.8	5.0
B12-C_4	1170	1184	1.2	48.3	49.6	2.6	27.8	29.1	4.5
B12-C_3	1217	1231	1.1	43.8	45.3	3.3	27.6	29.2	5.5
B13-C_4	836	839[*]	0.4	48.5	49.5[a#]	2.0	26.3	25.4[#]	3.5
B13-C_3	828	836[*]	1.0	44.4	44.9[a*]	1.1	26.0	25.9[*]	0.4
B14-C_4	849	839	1.2	49.6	49.4[b]	0.4	27.9	27.3	2.2
B14-C_3	831	837	0.7	45.2	44.7[a]	1.1	27.1	27.5	1.5
B15-C_4	795	806	1.4	51.2	52.2[a#]	1.9	28.3	27.8	1.8
B15-C_3	795	801	0.7	46.6	47.9[a]	2.7	28.3	27.4	3.3
B16-C_4	928	949	2.2	49.5	50.3[b]	1.6	27.9	29.3	4.8
B16-C_3	985	968	1.8	44.9	47.0[b]	4.5	28.5	29.8	4.4

Unlabeled experimental data are from ref [1].

[a]Data from ref [3].

[b]Data from ref [4].

[*]Data obtained by averaging properties of two isomers.

[#]Data obtained by fitting the properties of corresponding homologues.

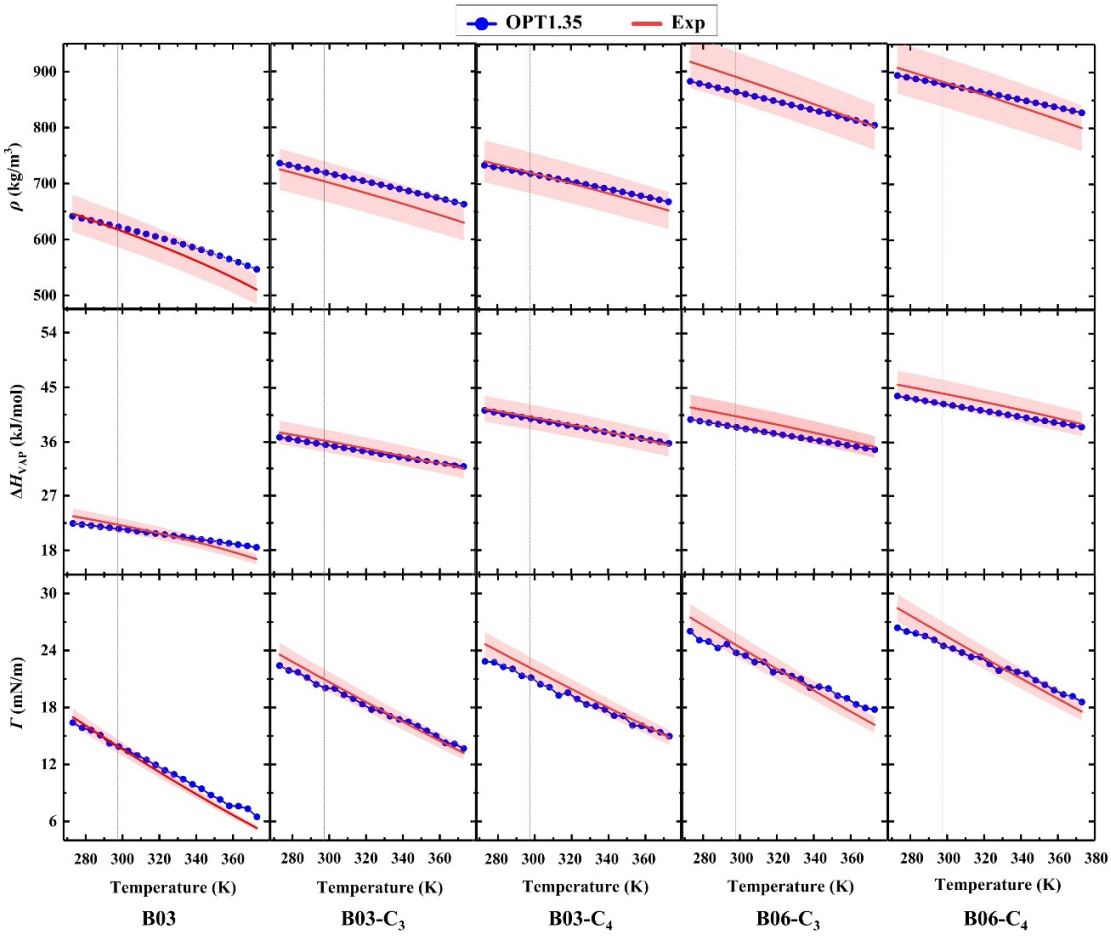

Figure S2 Thermodynamic properties ρ, ΔH_{VAP} and Γ of the homologs of B03 and B06 at temperatures from 273 K to 373 K. The experimental data were derived from ref [1] and ΔH_{VAP} were corrected at 298 K [3]. The shadow indicates the deviation range of ±5% from experimental data.

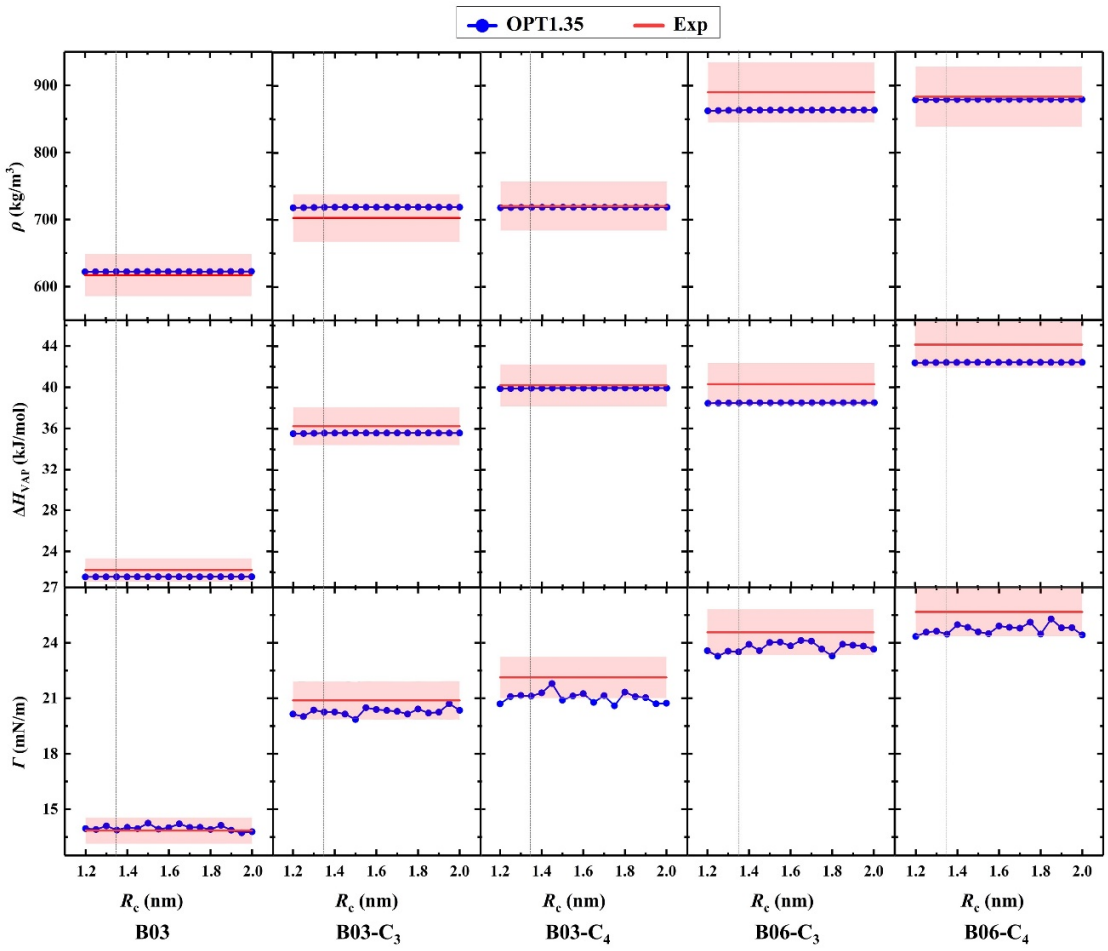

Figure S3 Thermodynamic properties ρ, ΔH_{VAP} and Γ of the homologs of B03 and B06 at truncation distance R_c from 1.20 nm to 2.00 nm. The shadow indicates the deviation range of ±5% from experimental data [1,3].

Table S6 Density ρ and heat of vaporization ΔH_{VAP} of fatty acid methyl esters calculated from simulations by using the OPT1.35 FF, as well as measured in experiment (Exp).

Molecules	CG-Names	ρ (kg/m^3)		ΔH_{VAP} (kJ/mol)	
		Sim	Exp[a]	Sim	Exp[b]
methyl-heptanoate	B06-C_3-C_3	848	877	54.6	51.8
methyl-octanoate	B06-C_4-C_3	859	873	57.9	56.9
methyl-nonanoate	B06-C_4-C_4	864	871	62.4	61.6
methyl-decanoate	B06-C_3-C_3-C_3	839	869	70.5	66.1
methyl-undecanoate	B06-C_4-C_3-C_3	848	867	74.1	70.8
methyl-dodecanoate	B06-C_4-C_4-C_3	854	869	77.8	76.6
methyl-tridecanoate	B06-C_4-C_4-C_4	860	859	82.9	80.0
methyl-tetradecanoate	B06-C_4-C_3-C_3-C_3	842	863	90.4	85.9
methyl-pentadecanoate	B06-C_4-C_4-C_3-C_3	848	862	94.3	89.3
methyl-linoleate	B06-C_3-C_3-B03-B03-C_3	881	889	108.3	102.2
methyl-oleate	B06-C_3-C_3-B03-C_4-C_3	860	870	108.5	103.3

[a]Experimental ρ from ref [1].

[b]Experimental ΔH_{VAP} from ref [4].

Calculated data in black, blue, and green denote that they have deviations from experiment in the range of $\delta < 2.5\%$, $2.5\% \le \delta < 5.0\%$, and $5.0\% \le \delta \le 7.0\%$, respectively.

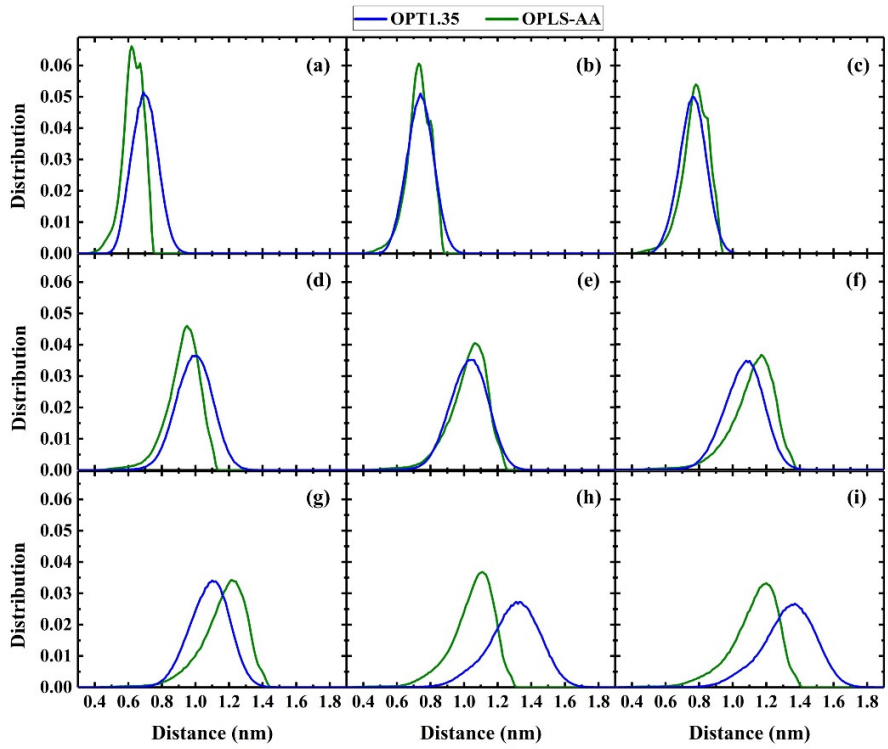

Figure S4 End-to-end distance distribution for methyl-heptanoate (a), methyl-octanoate (b), methyl-nonanoate (c), methyl-decanoate (d), methyl-undecanoate (e), methyl-dodecanoate (f), methyl-tridecanoate (g), methyl-tetradecanoate (h), and methyl-pentadecanoate(i) single chains obtained from CG and AA-mapped trajectories, respectively.

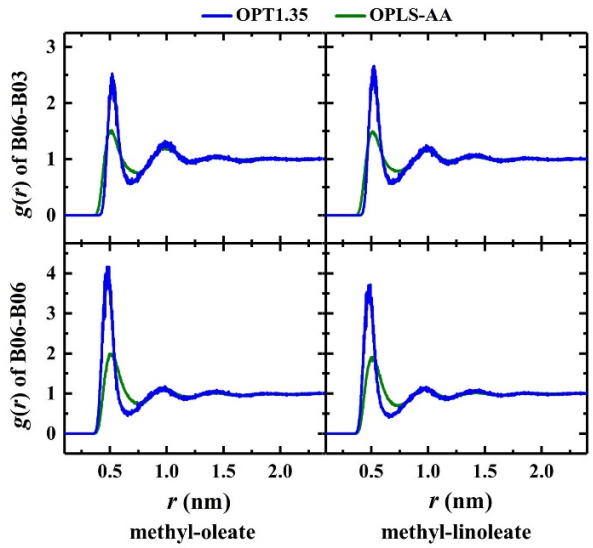

Figure S5 Radial distribution functions $g(r)$ of intermolecular B06~B03 and B06~B06 pairs in methyl-oleate and methyl-linoleate bulks.

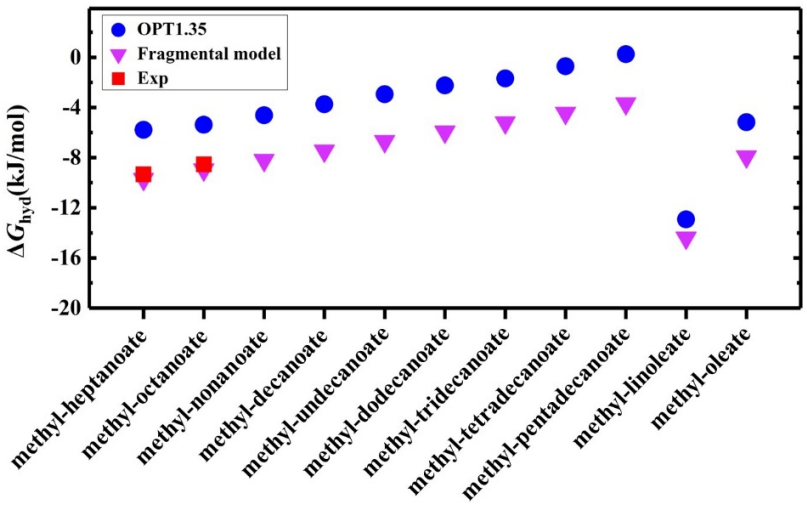

Figure S6 Hydration free energy ΔG_{hyd} of FAMEs. Data predicted by OPT 1.35 FF were compared to those estimated by fragmental model [5]. The experimental data for methyl-heptanoate and methyl-octanoate are from ref [6].

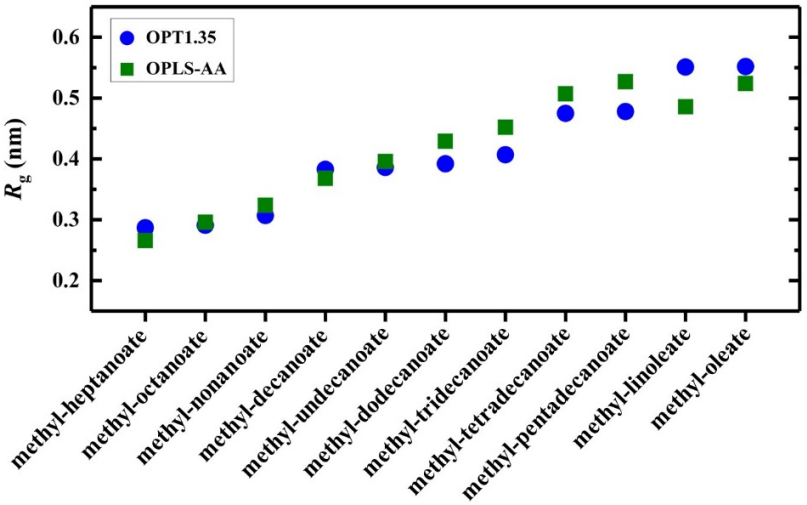

Figure S7 Radius of gyration of FAMEs in water.

Reference

1. C. L. Yaws, *Thermophysical properties of chemicals and hydrocarbons*. 2 ed.; William Andrew: Beaumont, **2014**.
2. C. L. Yaws, *The Yaws handbook of vapor pressure: Antoine coefficients*. 2 ed.; Gulf Professional Publishing: Houston, **2015**.
3. J. Rumble, *CRC Handbook of Chemistry and Physics*. 97 ed.; CRC press: Boca Raton, **2017**.
4. J. S. Chickos, W. E. A., Enthalpies of Vaporization of Organic and Organometallic Compounds, 1880–2002. *J. Phys. Chem. Ref. Data* **2002,** *32*, 519.
5. Pepe, G.; Guiliani, G.; Loustalet, S.; Halfon, P., Hydration free energy a fragmental model and drug design. *European Journal of Medicinal Chemistry* **2002**, 37 (11), 865-872.
6. A. V. Marenich, C. P. K., J. D. Thompson, G. D. Hawkins, C. C. Chambers, D. J. Giesen, P. Winget, C. J. Cramer, D. G. Truhlar Minnesota Solvation Database – version 2012, University of Minnesota, Minneapolis, **2012**.