

Supporting Information to Liquid Ethylene Glycol: Prediction of Physical Properties, Conformer Population and Interfacial Enrichment with a Refined Non-Polarizable Force Field

Anjali Gaur and Sundaram Balasubramanian*

Chemistry and Physics of Materials Unit

Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064, India

E-mail: bala@jncasr.ac.in

Contents

S1	Atom Type Nomenclature	S3
S2	Nomenclature of Ethylene Glycol Conformations	S3
S3	Quantum Calculations of Ethylene Glycol Monomer	S4
S4	EG Dimer Potential Energy Scan	S6
S5	Details of Ab Initio Molecular Dynamics Simulation	S7
S6	Force Field Parametrization	S9

S7	Structure of Liquid EG	S10
S7.1	Radial Distribution Functions	S10
S7.2	Structure Factor	S13
S8	Details of Force Field Parameters and Classical MD Simulation	S15
S8.1	Our Refined Force Fields: FF-v1 and FF-v2	S15
S8.2	Model 1	S19
S8.3	Model 2	S21
S8.4	Model 3	S21
S8.5	Model 4	S22
S9	Well-tempered Metadynamics Simulation	S23
S10	Simulation Details and Results For Physical Properties	S27
S10.1	Density	S27
S10.2	Self-diffusion	S28
S10.3	Shear Viscosity	S31
S10.4	Heat of Vaporization	S34
S10.5	Surface Tension	S35
S10.6	Static Dielectric Constant	S37
S11	Ethylene Glycol Orientation at Liquid-vapor Interface	S39
	References	S40

S1 Atom Type Nomenclature

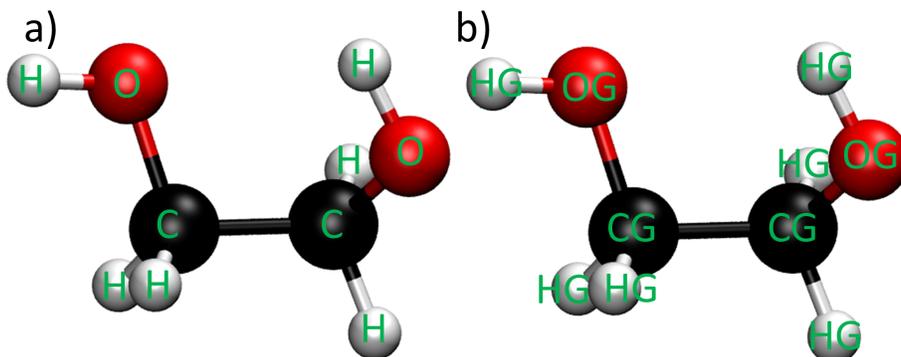


Figure S1: Structure of ethylene glycol molecule a) labeled with atom names b) labeled with atom types used in the force field.

S2 Nomenclature of Ethylene Glycol Conformations

The nomenclature of ethylene glycol (EG) conformations are based on dihedral angle values of its three dihedral angles H-O-C-C (terminal), O-C-C-O (central), and C-C-O-H (terminal) and shown in Figure S2. When a dihedral angle is greater than or equal to 150° or less than or equal to -150° , it is termed as *trans* and when it is between $\pm 30^\circ$ to $\pm 90^\circ$ it is termed as *gauche*. Symbol T and G are used if the OCCO dihedral is in *trans* and *gauche* conformation, respectively. Lower case symbols t (*trans*) and g (*gauche*) are used for clockwise rotation of terminal dihedrals and t' (*trans*) and g' (*gauche*) for anticlockwise rotation of the terminal dihedrals. For example, figure S2a is named as gTg' and figure S2b is named as tGg'. In common discussion, a EG molecule is referred to be in *trans* or *gauche* conformation based on the dihedral angle of the OCCO dihedral.

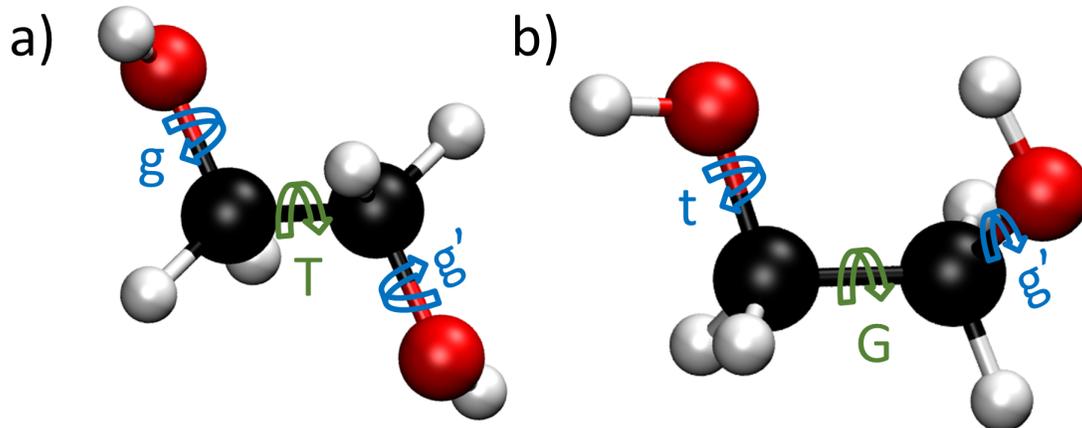


Figure S2: Structure of EG molecule shown in a) gTg' conformation and b) tGg' conformation.

S3 Quantum Calculations of Ethylene Glycol Monomer

Ab initio geometry optimization calculations in gas-phase and implicit EG solvent medium were performed for a single EG molecule using Gaussian 16¹ software. EG molecules exist in ten unique conformations. The ten conformers can be divided into two categories, one in which the OCCO dihedral angle adopts *trans* conformation and the other wherein it adopts *gauche* conformation. We performed calculations for the two most stable conformers in the respective two categories, *viz.* gTg' and tGg'. GaussView² was used to build the initial structures. Density functional theory (DFT) along with 6-311++G(d,p) basis set and B3LYP exchange and correlation functional^{3,4} were used for geometry optimization. Dispersion interactions were corrected using Grimme's D3⁵ correction along with Becke and Johnson's (BJ)⁶ damping function. A convergence criterion of 10^{-9} a.u. and 10^{-7} a.u. was used on the density matrix and energy matrix, respectively. And using the "opt=verytight" keyword within the Gaussian 16 software, a convergence criterion of 2×10^{-6} a.u. on the maximum force and 10^{-6} a.u. on the root-mean-square force was employed. Ultrafine integration grid was used for numerical integration. No symmetry constraints were applied for geometry optimization. Frequency calculations were performed on the optimized geometries to confirm energy minima. The energy difference between gTg' and tGg' conformers and their dipole

moment values are presented in Table S1.

Table S1: Potential energy difference in gas phase and implicit EG solvent medium between two conformers of EG. Energies and dipole moments were calculated using DFT at B3LYP-D3BJ/6-311++G(d,p) level of theory.

	E(gTg') - E(tGg') (kJ/mol)	Dipole Moment (Debye)	
		gTg'	tGg'
Gas-phase	11.61	0.00	2.63
Implicit solvent (PCM)	7.11	0.00	3.21

Geometry optimization of the two, gTg' and tGg', conformers were also performed using CP2K software. DFT with molecularly optimized double- ζ (DZVP-MOLOPT-SR-GTH) basis set and Perdew, Burke, and Ernzerhof (PBE)^{3,7-9} functional with corresponding Goedecker, Teter, and Hutter (GTH)⁸⁻¹⁰ pseudopotential was used for geometry optimization calculation. Dispersion interactions were corrected using Grimme's D2 dispersion correction.¹¹ The gas phase calculations were performed at this level of theory because *ab initio* molecular dynamics (AIMD) simulations of liquid EG were performed at the same level, and later we will compare the dipole moment of the single molecule with dipole moment distribution obtained from the AIMD simulations. A density cutoff of 700 Ry was used. A target accuracy of 10^{-7} a.u. was employed for SCF convergence. The energy difference between gTg' and tGg' conformers in the gas phase and their dipole moment values are present in Table S2. They demonstrate the close agreement of the molecular energetics and electrostatics represented by PBE-D3 with that of the B3LYP-D3BJ calculations.

Table S2: Potential energy difference in gas phase between two conformers of EG. Energies and dipole moments were calculated using DFT at GTH, PBE-D2/DZVP-MOLOPT-SR level of theory.

	E(gTg') - E(tGg') (kJ/mol)	Dipole Moment (Debye)	
		gTg'	tGg'
Gas-phase	9.72	0.00	2.41

S4 EG Dimer Potential Energy Scan Along O-H...O Hydrogen Bond Direction

One of our main objectives to perform liquid phase AIMD simulations was to ascertain the equilibrium population of *trans* and *gauche* conformers; thus, it is crucial to choose functionals and protocols carefully so as to be as close to more accurate treatments of intermolecular interactions as possible. Given that periodic DFT (as implemented in CP2K, say) has an efficiency limitation in handling hybrid functionals, van der Waals interaction needed to be included through the empirical Grimme approach. To benchmark the same, we performed a quantum potential energy rigid scan of an EG dimer along the direction of the intermolecular hydrogen bond at different levels of commonly used DFT methods and CCSD(T) theory. In this calculation, both the EG molecules were in *gauche* conformation, as shown in Figure S3.

The scans were performed using two commonly used exchange correlation functionals BLYP^{3,4} and, PBE^{3,7-9} with corresponding GTH⁸⁻¹⁰ pseudopotential to represent core electrons.^{3,7-9} The molecularly optimized double- ζ (DZVP-MOLOPT-SR-GTH) basis set was employed to represent valence electrons.¹² Dispersion interactions for PBE functional were corrected using Grimme's D2 and D3 dispersion correction.¹¹ The potential energy surfaces (PES) thus obtained were compared with CCSD(T) results. The potential energy profiles are shown in Figure S4. The PBE-D2/DZVP-MOLOPT-SR level of theory compares

the best with the CCSD(T) profile; hence AIMD simulations were performed at the PBE-D2/DZVP-MOLOPT-SR level of theory.

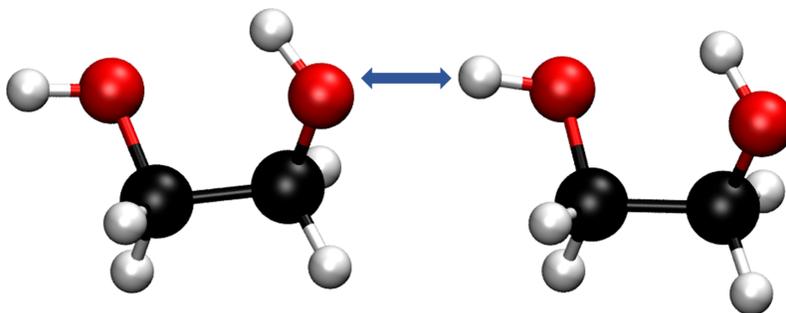


Figure S3: EG dimer, both the EG molecules are in *gauche* conformation. Blue arrow shows the direction of potential energy scan performed at different level of theories (see Figure S4). Color code: H: white, O: red, C: black.

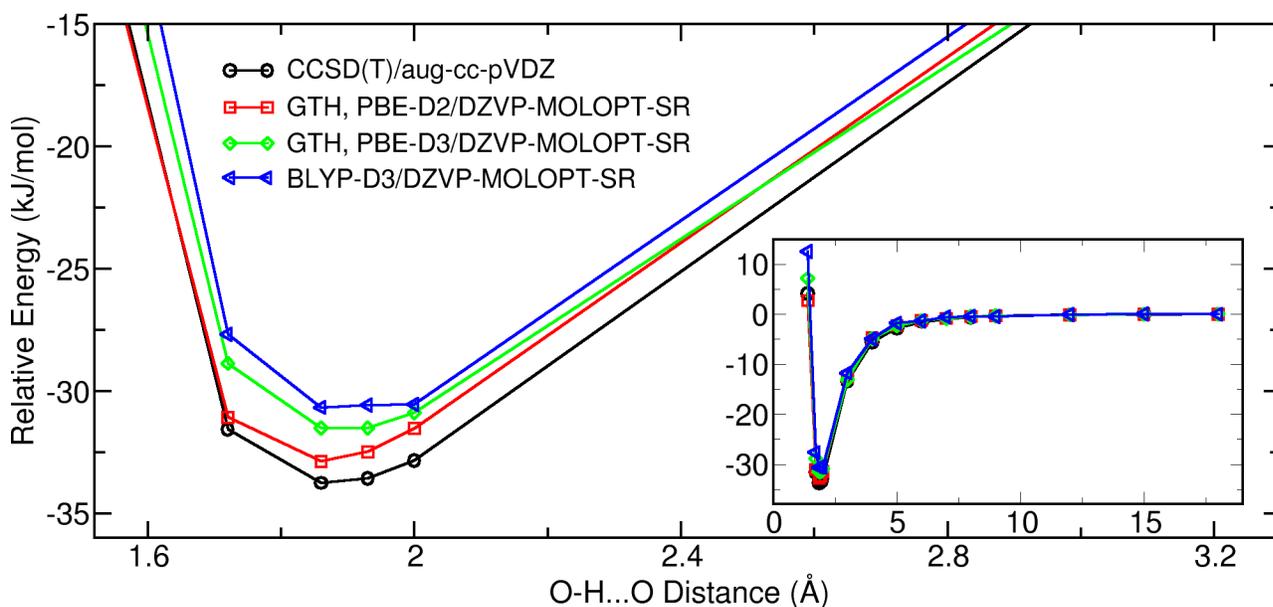


Figure S4: Potential energy of EG dimer along the O-H...O hydrogen bond scan direction, performed in gas phase, at different level of theories. Same data is presented both in the inset and in the main panel.

S5 Details of *Ab Initio* Molecular Dynamics Simulation

Jindal and Vasudevan had shown the existence of more than 20%¹³ conformers in the *trans* state. Yet, we felt the need for this result to be confirmed independently. Thus, we resorted

to two independent AIMD simulations of liquid EG, each initiated from a configuration that contained different fractions of *trans* conformers. These simulations of liquid EG were carried out using the CP2K¹⁴ software. A cubic box containing 100 EG molecules with 21.02 Å side length was constructed using PACKMOL,¹⁵ corresponding to the experimental density of 1.11 g/cc.

Table S3: System details used for EG AIMD simulations.

No. of Molecules	Cubic Box Length	No. of Valence Electrons
100	21.02 Å	2600

The initial configurations of liquid EG for the subsequent AIMD runs were obtained from a classical MD simulation trajectory carried out using the GROMACS-2018.3 package^{16,17} in the constant NVT ensemble for 10ns. This particular simulation was performed using the Model 1 force field (FF) with a refined OCCO dihedral potential which produced a mean *trans* population of 20%. Two snapshots, one with 15 % and another with 25 % of *trans* population were selected as initial configurations for the two independent AIMD runs. To perform the AIMD simulation, the level of the theory was chosen based on gas-phase scans of an EG dimer, as discussed in the previous section. The AIMD simulations were performed at the PBE-D2/DZVP-MOLOPT-SR level of theory. A density cutoff of 700 Ry was used. Periodic boundary conditions were applied in all three directions. A target accuracy of 10^{-7} a.u. was employed for SCF convergence. A NVT molecular dynamics simulation was performed using 0.5 fs time step for 40 ps. Temperature was controlled using Nosé-Hoover thermostat^{18,19} at 298.15 K.

All the results from AIMD simulations are calculated from the last 16 ps of the 40 ps trajectory and averaged over both the AIMD runs.

S6 Force Field Parametrization

As discussed in the introduction in the main manuscript, EG in liquid state has a *trans* population of 22 %. However, FFs in literature provide either nearly zero *trans* population or incorrect position of maximum of the OCCO dihedral distribution (see Table S12). Therefore we carried out FF parametrization to obtain the right fraction of *gauche* and *trans* population. Model 1 FF parameters (parameters provided in section S8.2) were used as starting point to obtain the refined FF parameters. Firstly, we performed a relaxed gas phase scan with respect to the central OCCO dihedral angle using quantum calculation at the B3LYP-D3BJ/6-311++G(d,p) level of theory. Using the geometries obtained after relaxed quantum scan, Model 1 FF parameters were used to calculate the relative energy of the conformers. Then we fitted the OCCO dihedral potential function to would match the relative energy of the conformers obtained from the quantum calculation (Figure S5). Thereafter, we performed a classical MD simulation of liquid EG using new OCCO dihedral parameters. The OCCO dihedral distribution thus obtained had a 20 % *trans* population. However, the density was 4.5% higher than the experimental density. In order to match the density, we compared the radial distribution function (RDF) of all possible atom type pairs with the RDFs computed using AIMD simulations to find clues for further refinement. Scaling ϵ of CG atom type by 0.5 produced a better comparison of density with experiments as well as of the RDFs with the AIMD RDFs. After every change in the FF parameter(s), we changed OCCO dihedral angle parameters manually to obtain a *trans* population of around 20 %. We calculated diffusion coefficient and viscosity at this stage and found that diffusion coefficient and viscosity were -71 % and 176 % of the experimental values, respectively. In order to improve on transport properties, we found scaling charges by 0.96 to give satisfactory results. After scaling the charges, we again looked at the central dihedral distribution and RDFs and changed OCCO dihedral parameters manually to obtain a *trans* population of around 20 %. The final refined FF produced a *trans* population of 20.9 %.

However, we observed that the HO-HO intramolecular RDF computed using refined

FF exhibited a shoulder at 2.2 Å which is not present in the RDF obtained from the AIMD simulation (Figure S7 right, green curve). This peak corresponds to the HO-HO distance when the EG molecule is in $g'Gg'$ conformation. we added a repulsive interaction between intramolecular HO-HO interactions to reduce the presence of $g'Gg'$ conformation and performed classical MD simulations. As shown in Figure S7 (Right, red curve), the shoulder at 2.2 Å disappeared. Henceforth, we call the refined FF without the intramolecular HO-HO repulsive interaction as FF-v1 and the FF with the 1-6 repulsive interaction as FF-v2.

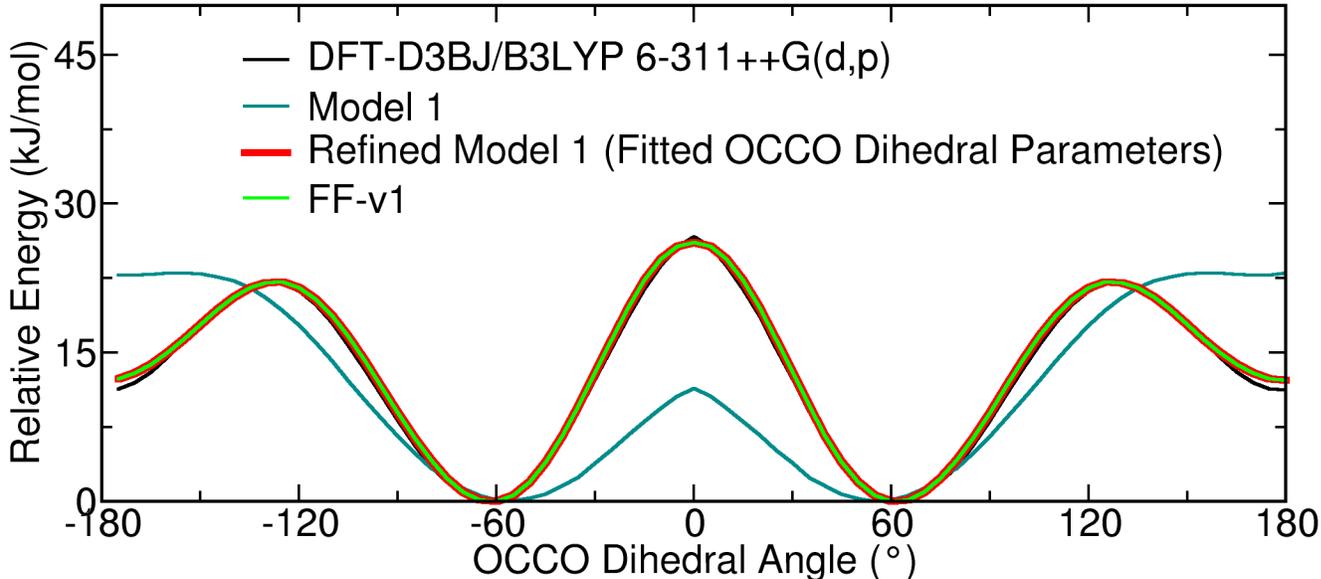


Figure S5: Gas phase scan along the OCCO dihedral angle.

S7 Structure of Liquid EG

S7.1 Radial Distribution Functions

Some of the radial distribution functions (RDFs) computed from the refined FF, FF-v1, and FF-v2 are shown in Figures S6, S7, and S8. They are in remarkable agreement with those computed from the AIMD simulations. The first peak of the OG-OG RDF obtained from the FF matches very well with that from AIMD, indicating that the FFs capture

hydrogen bonding accurately. All the RDFs obtained from both FF-v1 and FF-v2 FF are very similar except for the HO-HO intramolecular RDF. A shoulder at 2.2 Å seen in the RDF computed using FF-v1 FF vanishes after the addition of intramolecular HO-HO (1-6) repulsive interaction, and thus the $g(r)$ of FF-v2 is in better agreement with AIMD results. Such a 1-6 repulsive term is similar to the 1-5 repulsive term used in the TraPPE²⁰ FF for EG.

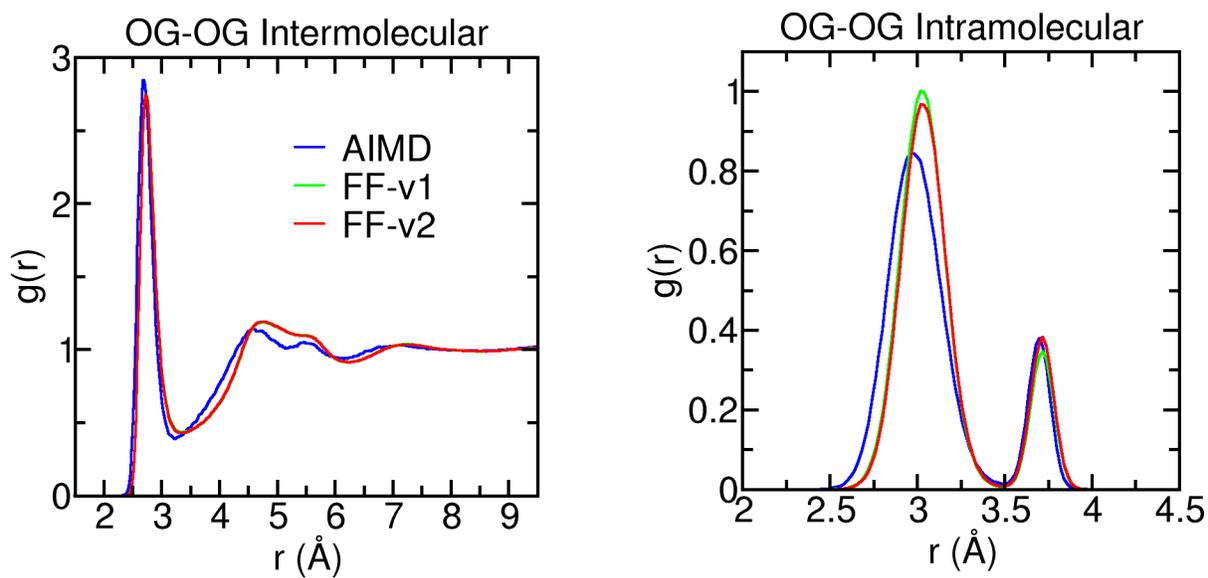


Figure S6: OG-OG intermolecular (left) and OG-OG intramolecular RDF (right) calculated from AIMD and classical force fields FF-v1 and FF-v2.

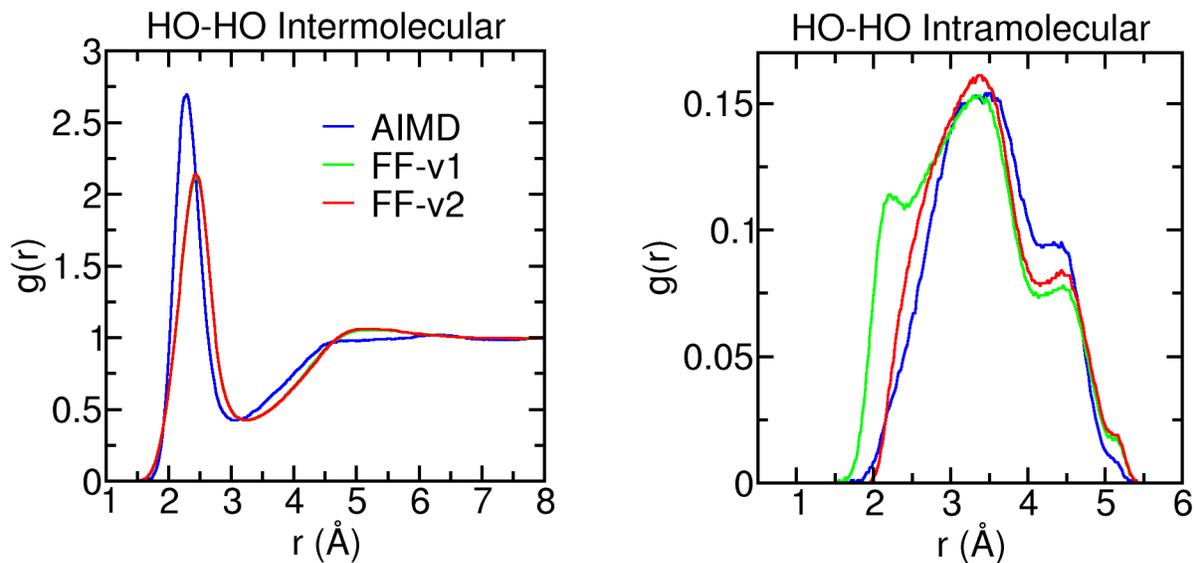


Figure S7: HO-HO intermolecular (left) and HO-HO intramolecular RDF (right) calculated from AIMD and classical force fields FF-v1 and FF-v2.

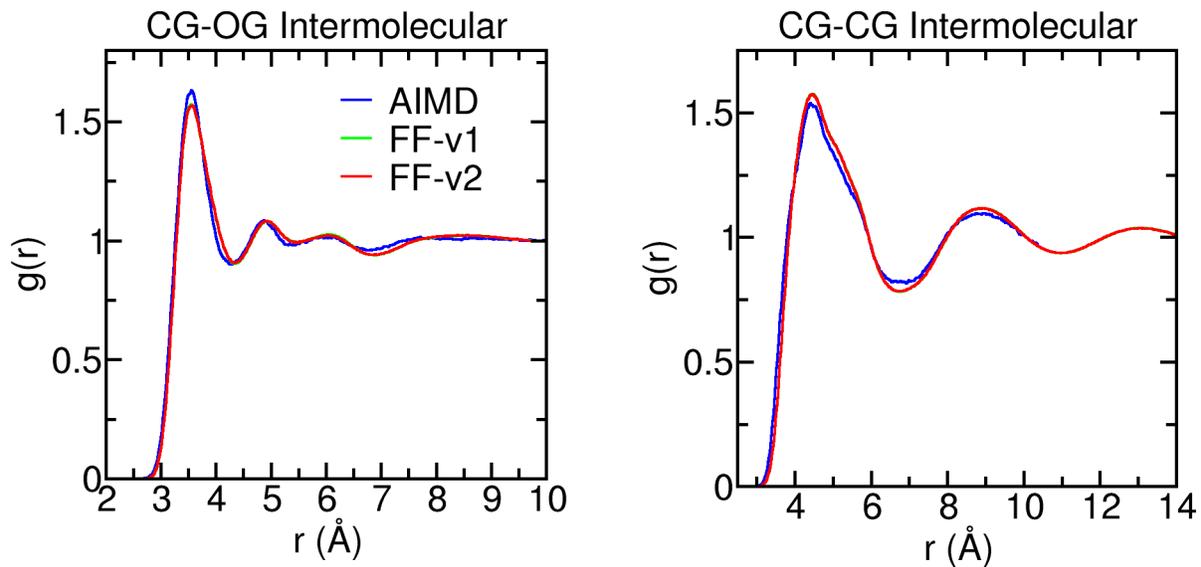


Figure S8: CG-OG intermolecular (left) and CG-CG intermolecular RDF (right) calculated from AIMD and classical force fields FF-v1 and FF-v2.

S7.2 Structure Factor

X-ray and neutron structure factor ($S(q)$) were computed using the following equation [S1](#).

$$S(q) = \frac{\sum_{\alpha} \sum_{\beta} x_{\alpha} x_{\beta} f_{\alpha}(q) f_{\beta}(q) 4\pi\rho \int_0^{\infty} r^2 (g_{\alpha\beta}(r) - 1) \frac{\sin qr}{qr} \omega(r) dr}{(\sum_{\alpha} x_{\alpha} f_{\alpha}(q))^2} \quad (\text{S1})$$

where x_{α} and x_{β} are fractions of atom types α and β . $f(q)$ is X-ray atomic form factor while calculating X-ray structure factor and neutron scattering length while calculating neutron structure factor. $g_{\alpha\beta}(r)$ is RDF between atom type α and β , $\omega(r)$ is a Lorch window function, $\omega(r) = \sin(\frac{2\pi r}{L}) / (\frac{2\pi r}{L})$, used to reduce noise at low wave-vectors. ρ is the number density and L is the length of the box. [Figure S9](#) shows X-ray and neutron structure factor computed using refined FF-v1 and FF-v2, and compared with experimental^{[21](#)} results.

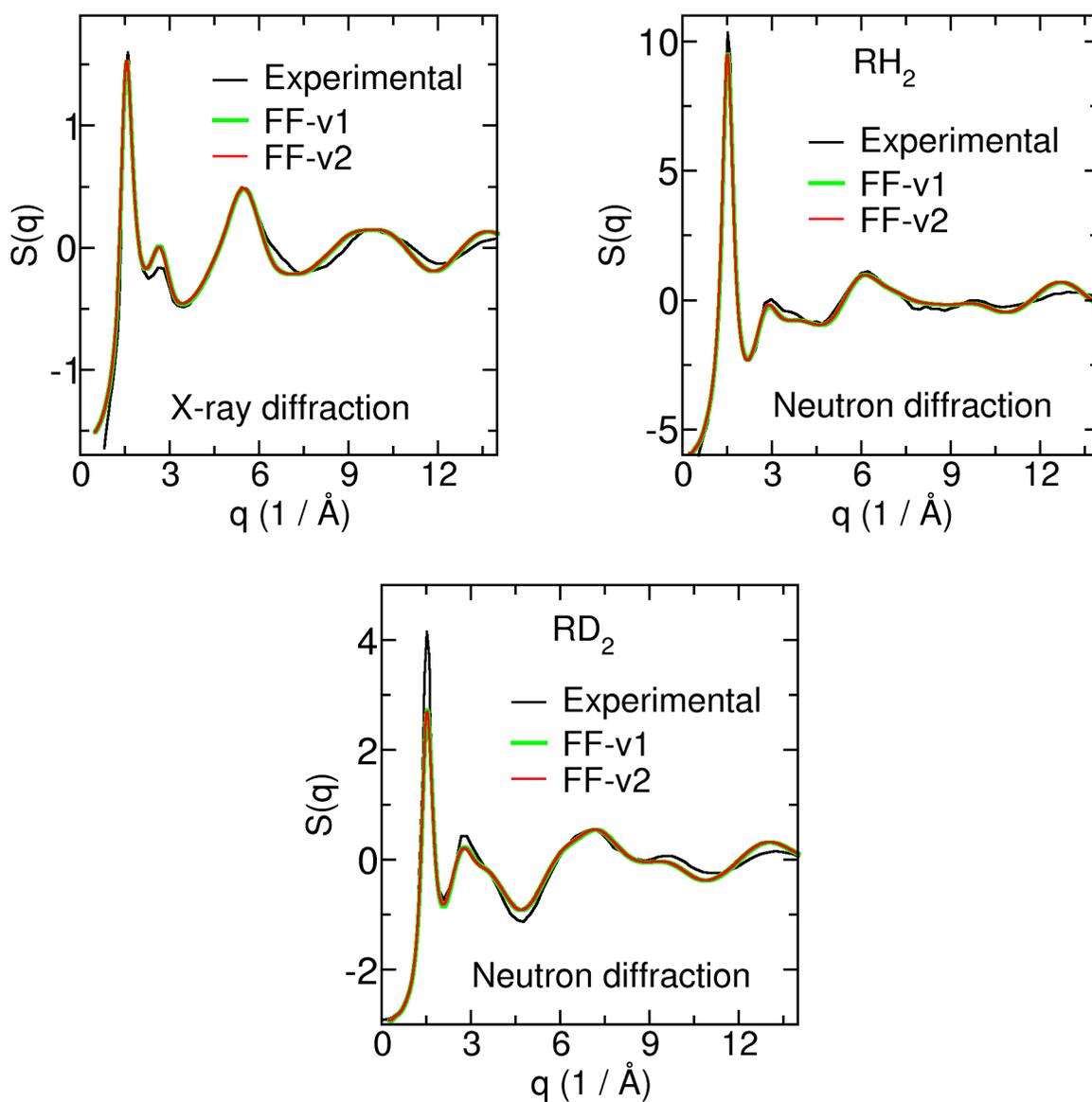


Figure S9: Structure factor for liquid EG calculated using our refined force fields, FF-v1 and FF-v2, and compared with experimental results. Experimental data is taken from Ref. 21. In the case of neutron diffraction, R denotes $\text{C}_2\text{D}_4\text{O}_2$, where the scattering length of deuterium is substituted in place of that of methylene hydrogen atoms.

S8 Details of Force Field Parameters and Classical MD Simulation

Following are the details of all the FFs that we have used. **Refined FF**: FF-v1 and FF-v2 are the FFs developed by us on top of Model 1 FF. Models 2, 3, and 4 were used to compute the OCCO dihedral distribution (Figure S10) to compare them with the distribution calculated from our refined FF. Simulations were performed using a cubic box, and periodic boundary conditions were applied in all three directions. Simulation box details for all the models are present in Table S12. We performed energy minimization followed by a 15 ns NPT run. Trajectories were saved every 1 ps. The last 5 ns of the NPT run were used to calculate average density and OCCO dihedral distribution. All the simulations were carried out using GROMACS-2018.3 software.^{16,17}

S8.1 Our Refined Force Fields: FF-v1 and FF-v2

Model 1 FF parameters (Section S8.2) were used as an initial point to obtain the refined FF parameters. FF-v1 is a general Amber FF (GAFF) based model and the functional form of the potential energy (E) is the same as that of GAFF. FF-v2 FF parameters are the same as FF-v1, except an extra 1-6 intramolecular repulsive interaction is included in FF-v2. The potential energy functional form for FF-v1 and FF-v2 is shown in Equations S2 and S3, respectively. Lennard-Jones (LJ) and Coulomb interactions are absent for atoms connected by less than three bonds. 1-4 LJ and Coulomb interactions are scaled by a factor of 0.5 and 0.833, respectively. A cutoff value of 10 Å was used to calculate short-range LJ and Coulomb interactions for all the simulations performed using FF-v1 and FF-v2 FF. FF-v1 and FF-v2 force field parameters are shown in Tables S4, S5, S6, and S7. The value of C used (equation

S3) is 40000 kJ mol⁻¹ nm¹².

$$\begin{aligned}
E_{FF-v1} &= E_{non-bonded} + E_{bonds} + E_{angles} + E_{dihedrals} \\
&= \sum_{i=1}^{N-1} \sum_{j>i}^N \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\} + \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angle} K_\theta (\theta - \theta_{eq})^2 \\
&+ \sum_{dihedrals} K_\phi (1 + \cos(n\phi - \phi_s))
\end{aligned} \tag{S2}$$

$$E_{FF-v2} = E_{FF-v1} + E_{repulsive}$$

$$E_{repulsive}(r_{ij}) = \sum_{ij} C / (r_{ij})^{12} \tag{S3}$$

where ij in $E_{repulsive}(r_{ij})$ are 1-6 intramolecular pairs.

Table S4: Non-bonded LJ parameters and charges used in FF-v1 and FF-v2 force field.

Atom Type	σ (nm)	ϵ (kJ/mol)	q (e)
CG	0.339967	0.228865	0.28669056
HG	0.247135	0.0656888	0.00033408
OG	0.306647	0.880314	-0.70818528
HO	0.000000	0.000000	0.42082656

Table S5: Bond parameters used in FF-v1 and FF-v2 force field.

Bond Type	K_r (kJ mol ⁻¹ nm ⁻²)	r_{eq} (nm)
CG-HG	276650	0.10969
CG-CG	251790	0.15375
CG-OG	265010	0.14233
OG-HO	310790	0.0985

Table S6: Angle parameters used in FF-v1 and FF-v2 force field.

Angle Type	K_θ (kJ mol ⁻¹ rad ⁻²)	θ_{eq} (degree)
CG-CG-HG	388.28	109.56
CG-CG-OG	564.84	110.19
CG-OG-HO	396.64	110.00
HG-CG-HG	328.03	108.46
HG-CG-OG	425.93	110.26

Table S7: Dihedral parameters used in FF-v1 and FF-v2 force field.

Dihedral Type	n	K_ϕ (kJ mol ⁻¹)	ϕ_s (degree)
CG-CG-OG-HO	3	0.20000	0.0000
	2	-0.60000	-180.0000
	1	1.00000	0.0000
HG-CG-CG-HG	3	0.65084	0.0000
HG-CG-CG-OG	1	1.04600	0.0000
HG-CG-OG-HO	3	0.69733	0.0000
OG-CG-CG-OG	3	6.30149	0.0453
	2	-5.20178	-179.9199
	1	7.81384	0.0481
	0	-3.94420	56.4662

A Leap-frog integrator was used to solve the equations of motion with a time step of 1 fs. A cutoff of 10 Å was used for short-range interactions, and energy and pressure corrections were applied. Long-range Coulomb interactions were calculated using the particle mesh Ewald method.²² Nosé-Hoover thermostat¹⁹ with 1 ps time constant was used to control temperature. Parrinello-Rahman barostat^{23,24} with a time constant of 10 ps was used for pressure coupling. Periodic boundary conditions were applied in all three directions. LINCS algorithm²⁵ was used to constraint all the covalent hydrogen bonds. To prepare the liquid EG box, 1400 EG molecules were first packed in a 50 Å box using PACKMOL,¹⁵ and then energy minimization was performed. Subsequently, a 25 ns NPT simulation was performed at various temperatures, and the average density was calculated using the last 5 ns. Several independent NVT simulations were performed to obtain other physical properties such as self-diffusion coefficient, shear viscosity, heat of vaporization, surface tension, and static dielectric constant.

Simulation run lengths and other calculation details related to each physical property are presented in section S10. Self-diffusion coefficient, shear viscosity, and static dielectric constant were calculated using in-house programs written in FORTRAN.

S8.2 Model 1

Model 1 FF was used as an initial framework to arrive at the refined FF parameters, FF-v1 and FF-v2. Model 1 is GAFF based model and its potential energy (E) functional form is the same as equation S2. Model 1 FF parameters were generated in the same way as reported in Reference 26. Bonded parameters and LJ parameters (σ and ϵ) were obtained using AmberTools,²⁷ and charges (q) were taken from reference 26. LJ and Coulomb interactions were absent for atoms connected by less than three bonds. 1-4 LJ and Coulomb interactions are scaled by a factor of 0.5 and 0.833, respectively. All the Model 1 FF parameters are presented in Table S8, S9, S10, and S11, for the sake of completeness. For all the simulations performed using this model, a cutoff value of 10 Å was used to calculate short-range LJ and Coulomb interactions.

Table S8: Non-bonded LJ parameters and charges used in Model 1 (literature S8.2) force field.

Atom Type	σ (nm)	ϵ (kJ/mol)	q (e)
CG	0.339967	0.457730	0.298636
HG	0.247135	0.0656888	0.000348
OG	0.306647	0.880314	-0.737693
HO	0.000000	0.000000	0.438361

Table S9: Bond parameters used in Model 1 force field.

Bond Type	K_r (kJ mol ⁻¹ nm ⁻²)	r_{eq} (nm)
CG-HG	276650	0.10969
CG-CG	251790	0.15375
CG-OG	265010	0.14233
OG-HO	310790	0.0973

Table S10: Angle parameters used in Model 1 force field.

Angle Type	K_θ (kJ mol ⁻¹ rad ⁻²)	θ_{eq} (degree)
CG-CG-HG	388.28	109.56
CG-CG-OG	564.84	110.19
CG-OG-HO	396.64	107.26
HG-CG-HG	328.03	108.46
HG-CG-OG	425.93	110.26

Table S11: Dihedral parameters used in Model 1 force field.

Dihedral Type	n	K_ϕ (kJ mol ⁻¹)	ϕ_s (degree)
CG-CG-OG-HO	3	0.66944	0.00
	1	1.04600	0.00
HG-CG-CG-HG	3	0.65084	0.00
HG-CG-CG-OG	1	1.04600	0.00
HG-CG-OG-HO	3	0.69733	0.00
OG-CG-CG-OG	3	0.60250	0.00
	2	4.91620	0.00

S8.3 Model 2

FF parameters for Model 2 were of Reference 28. This model is based on CHARMM general FF (CGenFF), refined by Kashyap and coworkers. Readers are directed to Reference 28 for its functional form and parameters. We used a cutoff value of 12 Å with a switching function from 10 Å to 12 Å. The cutoff values, and switching functions for short-range interactions are the same as mentioned in the above reference. Short-range LJ and Coulomb interaction were excluded for atoms connected by less than three bonds. We note that the OCCO dihedral distribution computed by us using this FF is in good agreement with that reported by Kashyap et al 28.

S8.4 Model 3

We refer to the TraPPE²⁰ FF for EG developed by Siepmann and coworkers as Model 3. All the parameters and potential energy functional form were taken from the official website <http://trappe.oit.umn.edu/> of the TraPPE FF. 1-4 Coulomb interactions were scaled by

a factor of 0.5, and 1-4 LJ interactions are excluded. A 1-5 short-range repulsive interaction was added. A cutoff radius of 14 Å was used to compute short-range interactions. All the above details are the same as recommended by Siepmann and coworkers.²⁰

S8.5 Model 4

The OPLS-AA based FF refined by Szeftczyk and Cordeiro²⁹ is referred to here as Model 4. Its potential energy functional form and parameters were taken from Ref. 29. Short-range LJ and Coulomb interactions were computed using a cutoff radius of 14 Å and 12 Å respectively. Short-range interactions were excluded for atoms bonded via less than 3 bonds. 1-4 LJ and Coulomb interactions were scaled by 0.5. All the above details are the same as recommended by Szeftczyk and Cordeiro.²⁹

Table S12: Simulation box details for MD runs carried out using various force field models. Average density values are calculated at 298.15 K and 1 bar. OCCO dihedral distribution is present in Figure S10, position of maximum of the distribution and *trans* population is present in this table.

Model	Box Length (Å)	No. of Molecules	ρ (kg m ⁻³)	OCCO dihedral distribution peak (°) (Fig. S10)	Trans (%)
FF-v1 (this work)	50	1400	1096.57 ± 4.30	69	20.87
FF-v2 (this work)	50	1400	1096.53 ± 4.12	70	22.91
Model 1 (GAFF+Ref. 26)	38	512	1170.14 ± 5.61	69	0.396
Model 2 ²⁸	30	300	1136.93 ± 7.16	0	12.65
Model 3 ²⁰	34	400	1122.14 ± 9.01	58	0.049
Model 4 ²⁹	33	390	1116.23 ± 7.72	44	0.01

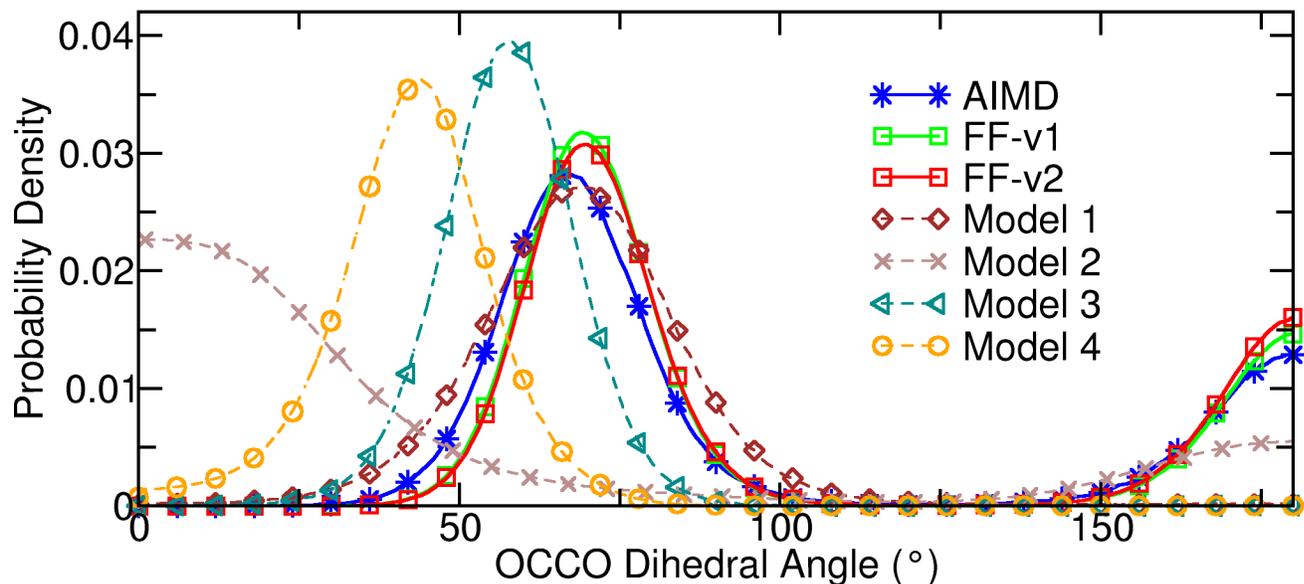


Figure S10: OCCO dihedral distribution obtained from classical MD simulations using our refined force fields (FF-v1 and FF-v2) and various other FFs compared against that obtained from AIMD simulations. Model 1 (GAFF + Ref. 26), 2,²⁸ 3²⁰ and 4²⁹ were taken from literature.

S9 Well-tempered Metadynamics Simulation

We carried out well-tempered metadynamics (WTMetaD)^{30,31} simulation and calculated the free energy difference between *trans* and *gauche* conformational states in liquid EG using our refined FFs FF-v1 and FF-v2. The WTMetaD simulation was performed using GROMACS-2018.3¹⁶ patched with PLUMED-2.6.2.^{32,33} We took an EG liquid snapshot from an equilibrated liquid NVT simulation at 298.15 K as the initial structure. OCCO dihedral angle of an EG molecule was chosen as the collective variable (CV), which was in *gauche* conformation (OCCO dihedral angle = 60.45 ° in case of FF-v1 and 74.88 ° in case of FF-v2). WTMetaD simulation parameters are presented in Table S13

The convergence of WTMetaD simulation was first checked qualitatively by observing the CV as a function of simulation time and the Gaussian hill height approaching a value close to zero (see Figure S11a, S11b). The convergence of free energy was confirmed by error

Table S13: Simulation parameters for WTMetaD simulation performed at 298.15 K.

Gaussian width	Initial Gaussian Height	Bias Deposition Rate	Bias factor
20.05 °	0.3 kJ/mol	0.5 ps	6

analysis using the block averaging method (Figure S11c). The free energy difference between *trans* and *gauche* conformation using FF-v1 FF is 1.7 kJ/mol (1.2 kJ/mol using FF-v2 FF), which amounts to an equilibrium *trans* population of 20.11 % (23.55 % using FF-v2 FF).

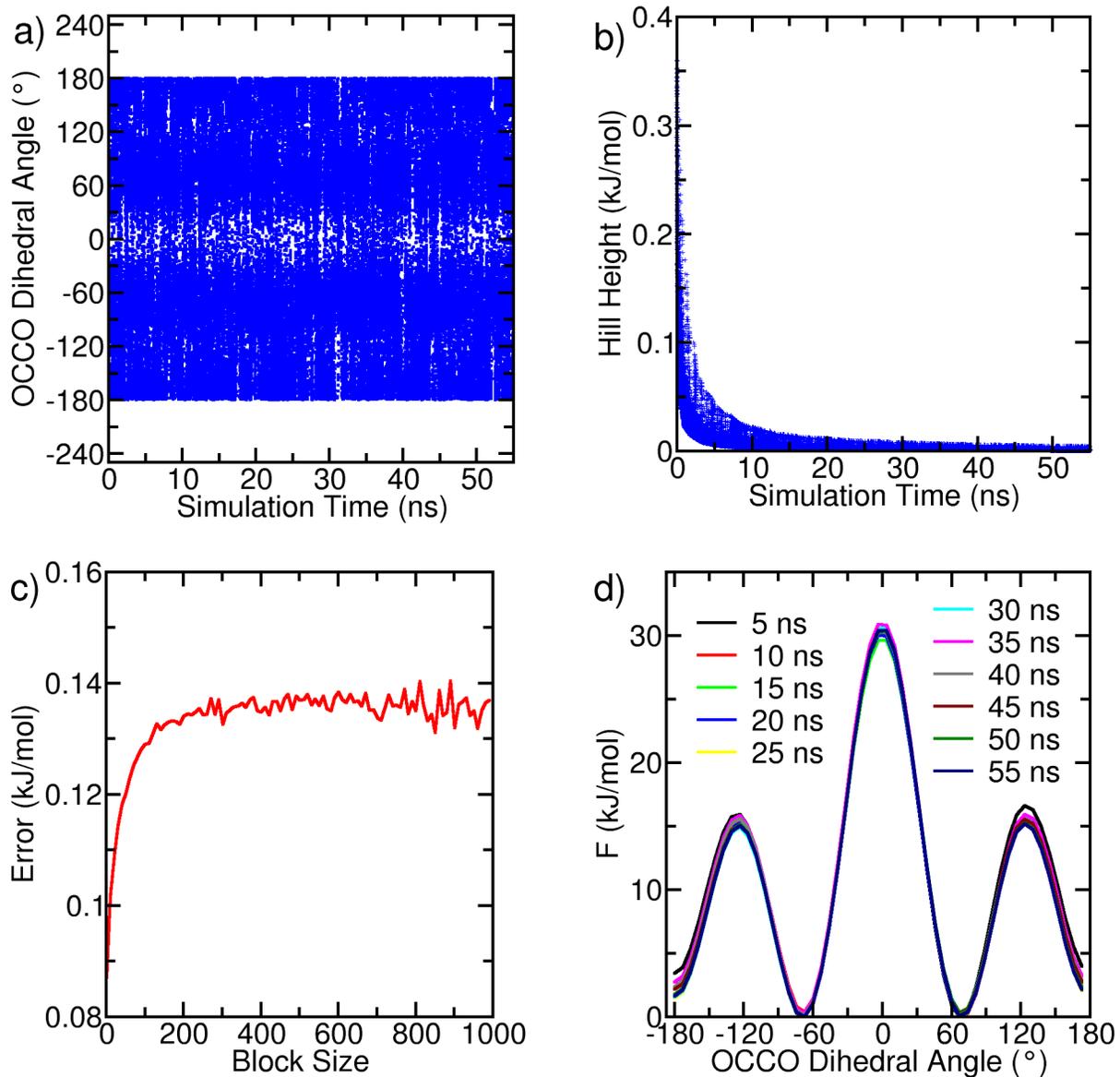


Figure S11: Using FF-v1 FF: (a) Time evolution of CV (OCCO dihedral angle) during 55 ns of WTMetaD simulation. (b) Gaussian hill height as a function of simulation time. The hill height approached nearly zero value within 40 ns of WTMetaD simulation. (c) Error analysis using different length of block size showing convergence of free energy. Maximum error in free energy is 0.15 kJ/mol. (d) Estimation of free energy as a function of OCCO dihedral angle calculated at every 5 ns.

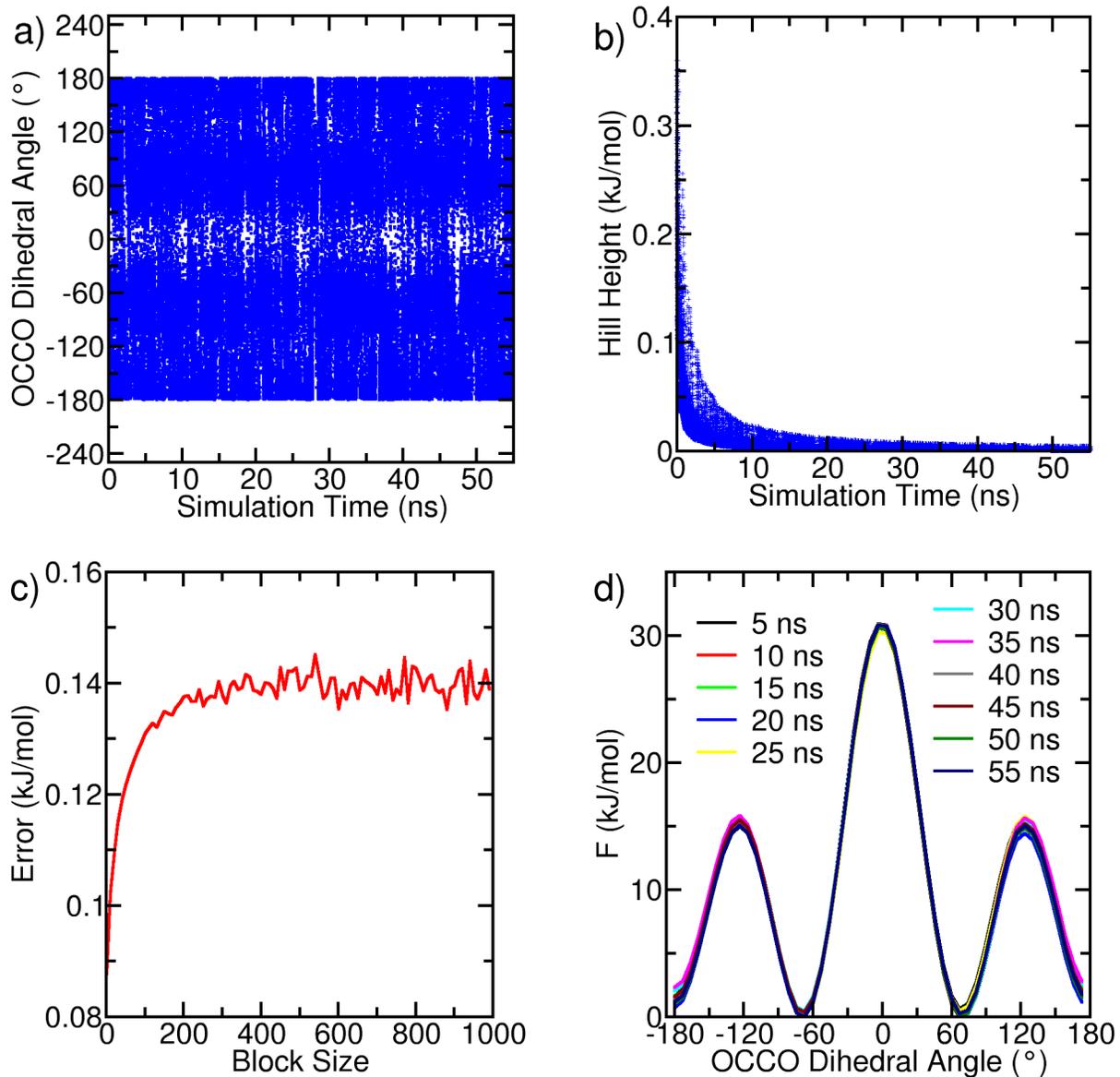


Figure S12: Using FF-v2 FF: (a) Time evolution of CV (OCCO dihedral angle) during 55 ns of WTMetaD simulation. (b) Gaussian hill height as a function of simulation time. The hill height approached nearly zero value within 40 ns of WTMetaD simulation. (c) Error analysis using different length of block size showing convergence of free energy. Maximum error in free energy is 0.14 kJ/mol. (d) Estimation of free energy as a function of OCCO dihedral angle calculated at every 5 ns.

S10 Simulation Details and Results For Physical Properties

S10.1 Density

Table S14: Simulation details used to estimate density using FF-v1 and FF-v2 models.

State Point	No. of Molecules	NPT	Independent runs
283.15K 1bar	1400	25 ns	1
293.15K 1bar	1400	25 ns	1
303.15K 1bar	1400	25 ns	1
313.15K 1bar	1400	25 ns	1
323.15K 1bar	1400	25 ns	1
333.15K 1bar	1400	25 ns	1
343.15K 1bar	1400	25 ns	1

The liquid density at different temperatures was estimated from the last 5 ns of 25 ns NPT simulations. Computed density and the corresponding experimental values are reported in Table [S15](#).

Table S15: Density at different state points, pressure is 1 bar at all the temperatures. Experimental data is taken from reference 34. Densities computed at different temperatures using refined FF are within 2.2 % of the experimental values.

Temperature (K)	Expt. Density (kg/m ³)	Simu. Density (kg/m ³)		Deviation (%)	
		FF-v1	FF-v2	FF-v1	FF-v2
283.15	1120.23	1109.87	1110.65	-0.92	-0.86
293.15	1113.27	1100.60	1100.58	-1.14	-1.14
303.15	1106.25	1091.45	1091.81	-1.34	-1.31
313.15	1099.17	1082.03	1082.84	-1.56	-1.49
323.15	1092.02	1072.76	1073.76	-1.76	-1.67
333.15	1084.78	1063.29	1064.32	-1.98	-1.89
343.15	1077.42	1053.80	1054.92	-2.19	-2.09

S10.2 Self-diffusion

Table S16: Simulation run details used to estimate self-diffusion coefficient using the FF-v1 and FF-v2 models.

Temperature	No. of molecules	NVT	Independent runs
298K	1400	50 ns	10
308K	1400	50 ns	10
318K	1400	50 ns	10

The mean square displacement (MSD) of molecules was calculated from ten independent simulations. Each simulation was run for 50 ns, and the initial 5 ns were discarded. The coordinates were dumped every 1 ps. The self-diffusion coefficient (D_{self}) of EG was calculated

using the Einstein relation (equation S4) at three different temperatures.

$$D_{self} = \frac{1}{6} \lim_{t \rightarrow \infty} \frac{d}{dt} \left\langle \frac{1}{N} \sum_{i=1}^N (\mathbf{r}_i(t) - \mathbf{r}_i(0))^2 \right\rangle \quad (\text{S4})$$

Where t is time, N is the total number of EG molecules, \mathbf{r} is the center of mass position of the EG molecules. MSD averaged over the ten independent runs is shown in Figure S13 (left). β , the diffusion exponent, was calculated using equation S5. β determines the regime of dynamics of the system; when β is equal to unity, the system is said to be in the diffusive regime.

$$\beta(t) = \frac{d \ln(MSD(t))}{d \ln(t)} \quad (\text{S5})$$

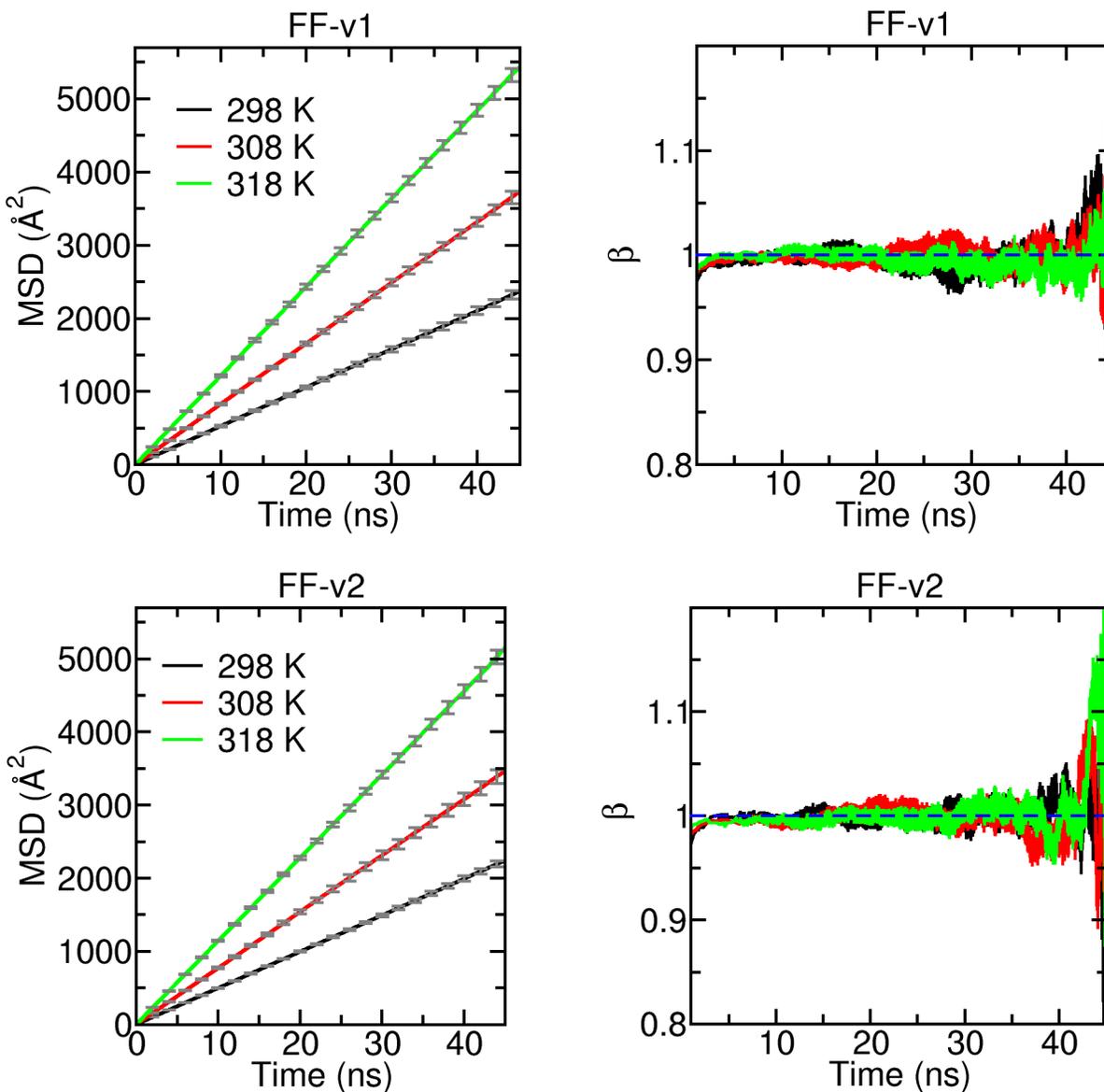


Figure S13: Mean square displacement (MSD) as function of simulation time at three different temperatures (left). The error bars represent standard deviation on mean, which is calculated from ten independent runs. MSD plots were used to compute self-diffusion coefficient. β was calculated using equation S5 to define the diffusive regime (right).

Table S17: Self-diffusion coefficient (D_{self}) values computed using refined FF and the average β value in the time interval in which D_{self} is computed. Experimental data is taken from Ref. 35. The uncertainty in computed D_{self} is calculated using block averaging method (6 blocks, each of 5 ns duration).

FF-v1						
Temperature (K)	Time Interval (ns)		Average β	D_{self} (10^{-7} cm ² /s)		ΔD (%)
	t_{start}	t_{end}		Experimental	This work	
298	5	35	0.994	8.34	8.74 ± 0.07	4.8
308	5	35	1.000	13.50	13.81 ± 0.07	2.3
318	5	35	0.997	19.00	20.18 ± 0.11	6.2
FF-v2						
Temperature (K)	Time Interval (ns)		Average β	D_{self} (10^{-7} cm ² /s)		ΔD (%)
	t_{start}	t_{end}		Experimental	This work	
298	5	35	0.999	8.34	8.29 ± 0.01	-0.6
308	5	35	0.999	13.50	12.84 ± 0.06	-4.9
318	5	35	0.999	19.00	18.97 ± 0.08	-0.2

S10.3 Shear Viscosity

Table S18: Simulation run details used to estimate shear viscosity using FF-v1 and FF-v2 models.

Temperature	No. of molecules	NVT	Independent runs
298.15K	1400	10 ns	40
313.15K	1400	10 ns	40
333.15K	1400	10 ns	40

Shear viscosity was calculated at three different temperatures. 40 independent simulation runs for 10 ns were performed at each temperature. All the nine components of the pressure tensor were dumped at every 1 fs. Traceless symmetric tensor at time t , $\tilde{\mathbf{P}}(t)$, was calculated to estimate viscosity using the Green-Kubo relation (Equation S6).

$$\eta = \frac{V}{10k_{\text{B}}T} \int_0^\infty \langle \tilde{\mathbf{P}}(t) : \tilde{\mathbf{P}}(0) \rangle dt \quad (\text{S6})$$

Figure S14 shows mean running viscosity (red color) calculated at 298.15 K along with running viscosity computed from the 40 independent simulations (gray color). Mean running shear viscosity calculated at three different temperatures is shown in Figure S15. The error bars represent the standard deviation on the mean. The flat region from 200 ps to 500 ps was divided into three equal parts to calculate mean viscosity and standard deviation on the mean. Viscosity values thus obtained were compared with experimental values, as reported in Table S19. All the computed values are within 16.3 % of the experimental values.

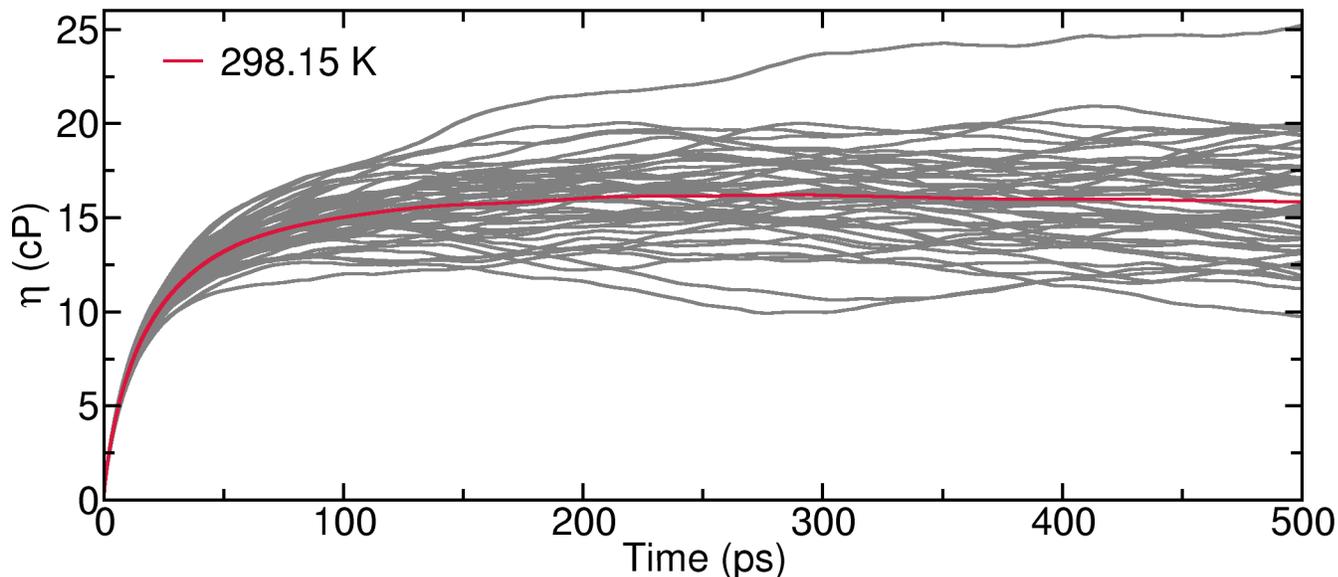


Figure S14: Gray lines represent running shear viscosity of 40 independent simulations and red line corresponds to mean running shear viscosity at 298.15 K, computed using refined FF-v1 force field. The mean shear viscosity converges smoothly as opposed to the running viscosity calculated using the individual independent simulations.

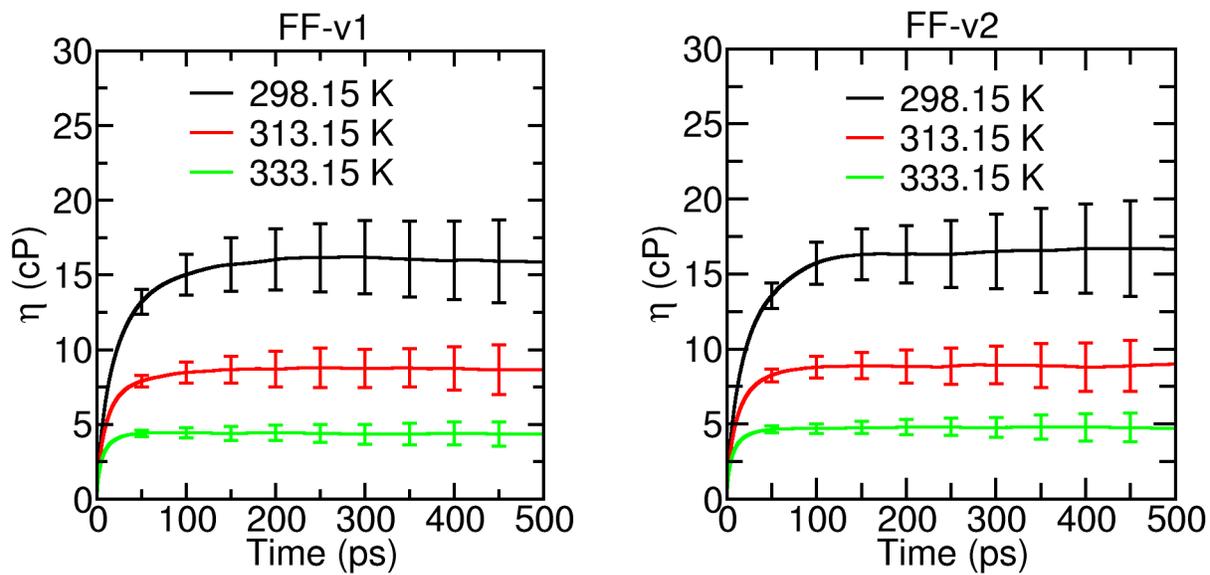


Figure S15: Mean running shear viscosity as function of simulation time at three different temperatures computed using refined FF.

Table S19: Shear viscosity values computed using refined FF. The computed values are compared with experimental values. Experimental data is taken from Ref. 36. The uncertainty in computed viscosity values is calculated using block averaging method (3 blocks, each of 100 ps duration).

FF-v1					
Temperature (K)	Time Interval (ps)		Viscosity (cP)		$\Delta\eta$ (%)
	t_{start}	t_{end}	Experimental	This work	
298.15	200	500	17.40	16.05 ± 0.09	-7.76
313.15	200	500	9.42	8.73 ± 0.04	-7.32
333.15	200	500	5.22	4.37 ± 0.01	-16.28
FF-v2					
Temperature (K)	Time Interval (ps)		Viscosity (cP)		$\Delta\eta$ (%)
	t_{start}	t_{end}	Experimental	This work	
298.15	200	500	17.40	16.54 ± 0.13	-4.94
313.15	200	500	9.42	8.88 ± 0.01	-5.73
333.15	200	500	5.22	4.78 ± 0.01	-8.43

S10.4 Heat of Vaporization

Heat of vaporization (ΔH_{vap}) was calculated at 298.15 K using three independent gas phase runs to compute the mean potential energy in the gas phase. The same for the liquid phase was calculated using a 20 ns NVT simulation. Assuming ideality in the gas phase and the same kinetic energy in both the gas and liquid phases, the formula used to calculate ΔH_{vap} is

$$\Delta H_{vap} = \langle E_g^{potential} \rangle - \langle E_l^{potential} \rangle + RT \quad (S7)$$

where $\langle E_g^{potential} \rangle$ and $\langle E_l^{potential} \rangle$ are the average potential energy in gas and liquid phase, respectively, R is the universal gas constant, and T is temperature.^{37,38} Computed values of ΔH_{vap} are present in the main manuscript.

S10.5 Surface Tension

Table S20: Details of MD simulations of liquid EG slab for the calculation of surface tension at 298.15 K using FF-v1 and FF-v2 model.

Box Length (Å)			No. of Molecules	Independent runs	NVT
L_x	L_y	L_z			
50.80	50.80	152.40	1400	4	5 ns

Slab MD simulations at 298.15 K were performed to calculate interface properties with box size $L_z = 3L_x = 3L_y$. In these simulations, a cutoff of 25 Å was used without long-range energy pressure correction because of the inhomogeneity of the system.³⁹ The slab MD simulation was run for 25 ns. The density profile, calculated using the last 5 ns, as a function of z is shown in Figure S16. The bulk liquid density of the slab simulation matches that obtained from the liquid NPT simulation. The density profile is fitted to Equation S8 with the typical hyperbolic tangent function.⁴⁰

$$\rho(z) = \frac{1}{2}(\rho_L + \rho_V) - \frac{1}{2}(\rho_L - \rho_V)\tanh[(z - z_0)/d] \quad (\text{S8})$$

where, ρ_L , ρ_V , z_0 , and d are density in liquid and vapor phase at liquid-vapor coexistence, the position of the Gibbs dividing surface, and the width of the liquid-vapor interface, respectively. The fitted density profile is shown with the dashed blue line in Figure S16. Fitting parameters ρ_L , ρ_V , z_0 , and d are provided in Table S21. ρ_L matches pretty well with the bulk liquid density obtained; ρ_L is just 0.13 % less than the bulk density.

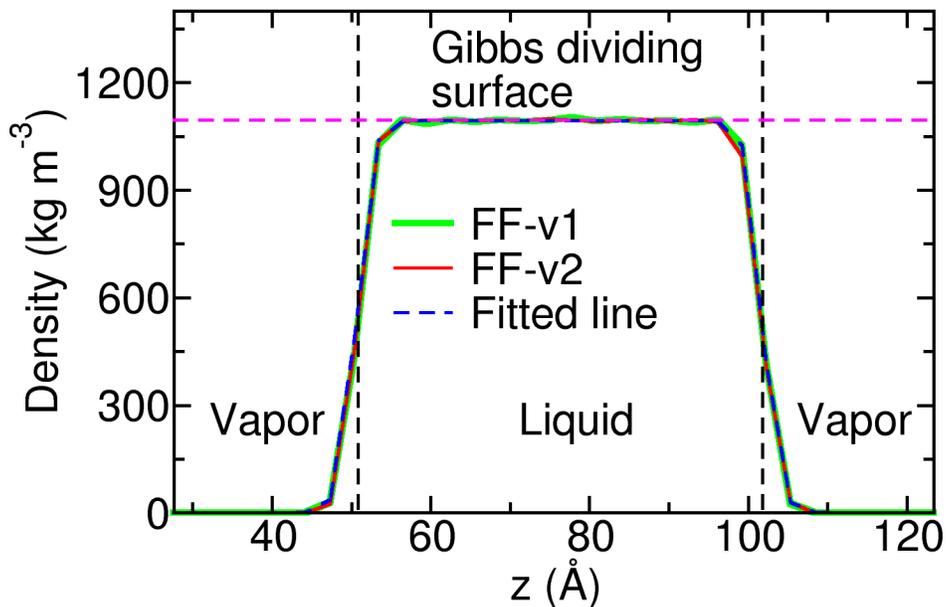


Figure S16: Density of EG as function of direction perpendicular to liquid-vapor interface (z -direction). Magenta dashed line is drawn at the bulk density obtained from liquid NPT simulation. Blue Dashed line is for fitted tangent hyperbolic function (Equation S8).

Table S21: Parameters obtained by fitting tangent hyperbolic function (Equation S8) to density profile.

ρ_L (kg m ⁻³)	ρ_V (kg m ⁻³)	z_0^L (Å)	z_0^R (Å)	d (Å)
1094.73	1.00	50.58	101.80	1.93

The last 20 ns of the NVT simulation was divided into four parts (each 5 ns). These four parts were considered independent simulations to calculate the surface tension. Diagonal pressure components (P_{xx}, P_{yy}, P_{zz}) were dumped every 1 fs, and each run was 5 ns long. Equation S9 was used to calculate the surface tension (γ). Surface tension values computed using FF-v1 and FF-v2 are presented in the main manuscript.

$$\gamma = \frac{L_z}{2} \left(\langle P_{zz} \rangle - \frac{\langle P_{xx} \rangle + \langle P_{yy} \rangle}{2} \right) \quad (\text{S9})$$

S10.6 Static Dielectric Constant

Table S22: Simulation run details used to estimate dielectric constant.

Model	Cubic Box Size	No. of Molecules	NVT (ns)	Independent runs
FF-v1	50.87 Å	1400	100	10
FF-v2	50.87 Å	1400	60	10

Ten independent simulations were performed to calculate the static dielectric constant. GROMACS-2018.3¹⁶ patched with PLUMED-2.6.2^{32,33} was used to write the total dipole moment of the simulation box (\vec{M}) every 1 fs. The static dielectric constant was calculated using fluctuations in \vec{M} ⁴¹ using equation S10.

$$\epsilon = 1 + \frac{4\pi}{3\epsilon_0 k_B T V} (\langle \vec{M}^2 \rangle - \langle \vec{M} \rangle^2) \quad (\text{S10})$$

Cumulative running averages of static dielectric constant calculated from the ten independent simulations using the FF-v1 FF are shown in Figure S17. Each simulation was run for 100 ns. The dielectric constant converges in 60 ns, hence to calculate the dielectric constant from the FF-v2 FF, we performed calculations for 60 ns rather than 100 ns (Figure S18).

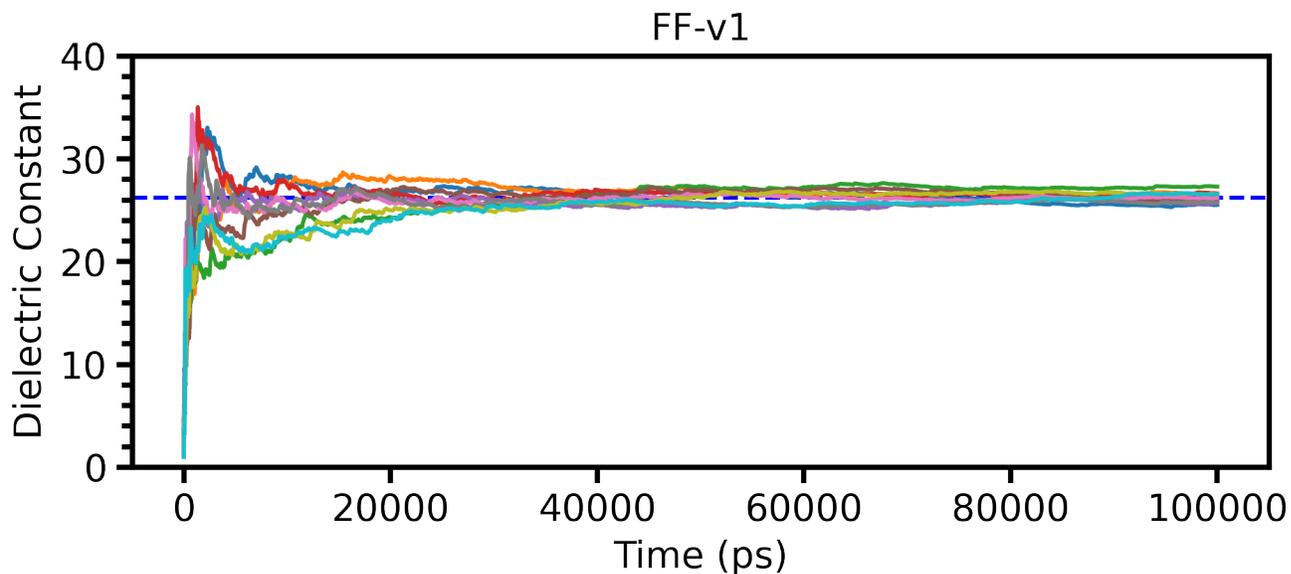


Figure S17: Cumulative running average of static dielectric constant from ten independent simulations using the FF-v1 force field. Each color corresponds to each independent simulation. The dashed blue line represents the mean static dielectric constant.

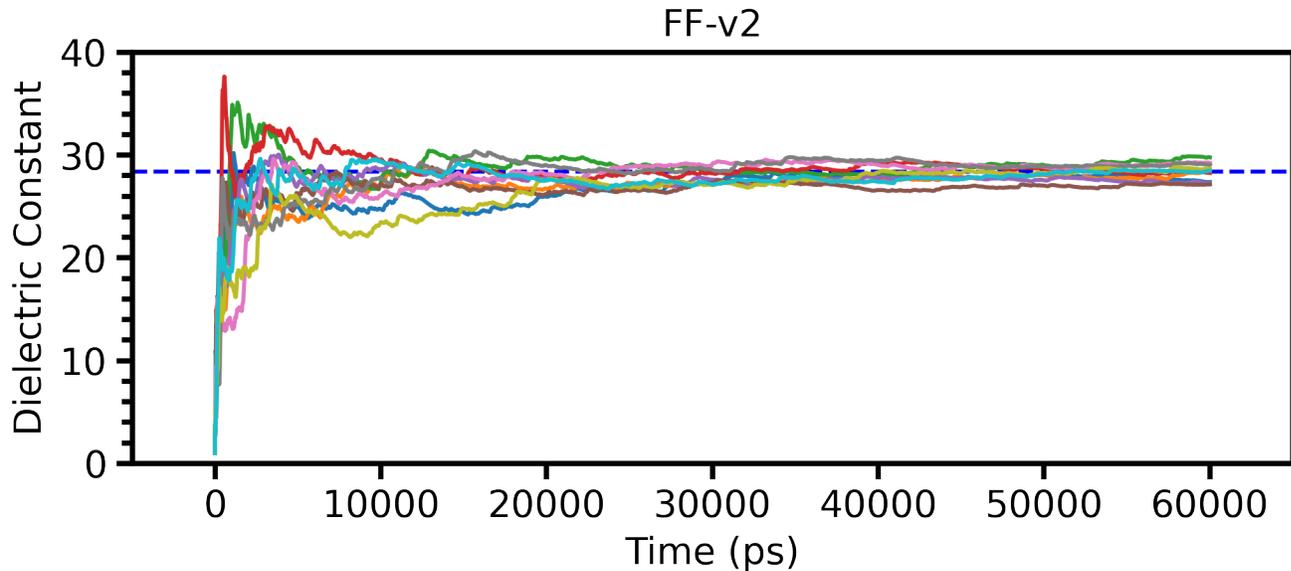


Figure S18: Cumulative running average of static dielectric constant from ten independent simulations using the FF-v2 force field. Each color corresponds to each independent simulation. The dashed blue line represents mean static dielectric constant.

S11 Ethylene Glycol Orientation at Liquid-vapor Interface

Liquid-vapor interface simulation details are the same as provided in section S10.5. EG orientation at the liquid-vapor interface was examined using four trajectories, each simulation being 5 ns long.

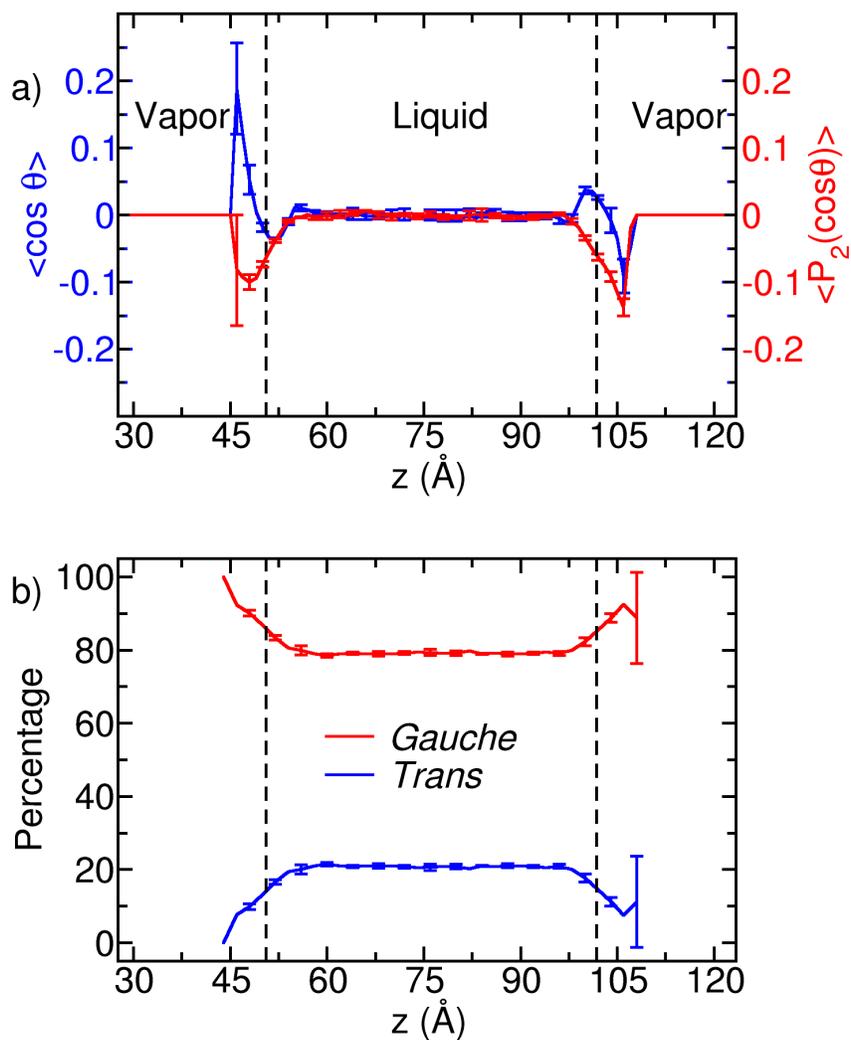


Figure S19: Plots computed using FF-v1 force field: a) $\langle \cos \theta \rangle$ and $\langle P_2(\cos \theta) \rangle$ along z-direction. b) Population percentage of *trans* and *gauche* conformation along z-direction. Error bars in both the figures represents standard deviation on mean. These are calculated by averaging over four trajectories each of 5 ns length.

References

- (1) Frisch, M. J. et al. Gaussian16 Revision A.03. 2016; Gaussian Inc. Wallingford CT.
- (2) Dennington, R.; Keith, T.; Millam, J. GaussView Version 5. Semichem Inc., Shawnee Mission, KS, 2009.
- (3) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (4) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.
- (5) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (6) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (7) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (8) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703–1710.
- (9) Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* **1998**, *58*, 3641–3662.
- (10) Krack, M. Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals. *Theor. Chem. Acc.* **2005**, *114*, 145–152.
- (11) Stefan, G. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

- (12) VandeVondele, J.; Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **2007**, *127*, 114105.
- (13) Jindal, A.; Vasudevan, S. Ethylene Glycol Dihedral Angle Dynamics: Relating Molecular Conformation to the Raman Spectrum of the Liquid. *J. Phys. Chem. B* **2021**, *125*, 1888–1895, PMID: 33560860.
- (14) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. CP2K: Atomistic Simulations of Condensed Matter Systems. *WIREs Comput. Mol. Sci.* **2014**, *4*, 15–25.
- (15) Martinez, L.; Andrade, R.; Birgin, E. G.; Martinez, J. M. Packmol: A Package for Building Initial Configurations for Molecular Dynamics Simulations. *J. Comput. Chem.* **2009**, *30*, 2157–2164.
- (16) Abraham, M. J.; van der Spoel, D.; Lindahl, E.; Hess, B.; the GROMACS development team, GROMACS User Manual. <http://www.gromacs.org>.
- (17) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447, PMID: 26620784.
- (18) Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511–519.
- (19) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255–268.
- (20) Stubbs, J. M.; Potoff, J. J.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 6. United-Atom Description for Ethers, Glycols, Ketones, and Aldehydes. *J. Phys. Chem. B* **2004**, *108*, 17596–17605.
- (21) Bakó, I.; Grósz, T.; Pálinkás, G.; Bellissent-Funel, M. C. Ethylene glycol dimers in

- the liquid phase: A study by x-ray and neutron diffraction. *J. Chem. Phys.* **2003**, *118*, 3215–3221.
- (22) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (23) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (24) Nosé, S.; Klein, M. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **1983**, *50*, 1055–1076.
- (25) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (26) Zhang, N.; Li, M.-R.; Zhang, F.-S. Structure and dynamics properties of liquid ethylene glycol from molecular dynamics simulations. *Chem. Phys. Lett.* **2019**, *718*, 12–21.
- (27) Case, D. et al. Amber 2018. 2018; <http://ambermd.org>.
- (28) Kaur, S.; Shobhna,; Kashyap, H. K. Insights Gained from Refined Force-Field for Pure and Aqueous Ethylene Glycol through Molecular Dynamics Simulations. *J. Phys. Chem. B* **2019**, *123*, 6543–6553, PMID: 31335141.
- (29) Szeferczyk, B.; D. S. Cordeiro, M. N. Physical Properties at the Base for the Development of an All-Atom Force Field for Ethylene Glycol. *J. Phys. Chem. B* **2011**, *115*, 3013–3019, PMID: 21388132.
- (30) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562–12566.
- (31) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, 020603.

- (32) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604–613.
- (33) Bonomi, M. et al. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670–673.
- (34) Tawfik, W. Y.; Teja, A. S. The densities of polyethylene glycols. *Chem. Eng. Sci* **1989**, *44*, 921–923.
- (35) Rodnikova, M. N.; Idiyatullin, Z. S.; Solonina, I. A. Mobility of molecules of liquid diols in the temperature range of 303–318 K. *Russ. J. Phys. Chem. A* **2014**, *88*, 1442–1444.
- (36) Ge, M.-L.; Ma, J.-L.; Chu, B. Densities and Viscosities of Propane-1,2,3-triol + Ethane-1,2-diol at $T = (298.15 \text{ to } 338.15) \text{ K}$. *J. Chem. Eng. Data* **2010**, *55*, 2649–2651.
- (37) Rizzo, R. C.; Jorgensen, W. L. OPLS All-Atom Model for Amines: Resolution of the Amine Hydration Problem. *J. Am. Chem. Soc.* **1999**, *121*, 4827–4836.
- (38) Wang, J.; Hou, T. Application of Molecular Dynamics Simulations in Molecular Property Prediction. 1. Density and Heat of Vaporization. *J. Chem. Theory Comput.* **2011**, *7*, 2151–2165, PMID: 21857814.
- (39) Zubillaga, R. A.; Labastida, A.; Cruz, B.; Martnez, J. C.; Snchez, E.; Alejandre, J. Surface Tension of Organic Liquids Using the OPLS/AA Force Field. *J. Chem. Theory Comput.* **2013**, *9*, 1611–1615, PMID: 26587622.
- (40) Alejandre, J.; Tildesley, D. J.; Chapela, G. A. Molecular dynamics simulation of the orthobaric densities and surface tension of water. *J. Chem. Phys.* **1995**, *102*, 4574–4583.
- (41) Gereben, O.; Pusztai, L. On the accurate calculation of the dielectric constant from molecular dynamics simulations: The case of SPC/E and SWM4-DP water. *Chem. Phys. Lett.* **2011**, *507*, 80–83.