Supporting Information for Towards Automated Benchmarking of Atomistic Forcefields: Neat Liquid Densities and Static Dielectric Constants from the ThermoML Data Archive

Kyle A. Beauchamp[+],[*,†] Julie M. Behr[+],[‡] Ariën S. Rustenburg,[¶] Christopher I. Bayly,[§] Kenneth Kroenlein,[‖] and John D. Chodera[*,†]

Computational Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, Tri-Institutional Program in Computational Biology and Medicine, Weill Cornell Medical College, New York, NY, Graduate Program in Physiology, Biophysics, and Systems Biology, Weill Cornell Medical College, New York, NY, OpenEye Scientific Software Inc., Santa Fe, NM, and Thermodynamics Research Center, NIST, Boulder, CO

E-mail: kyle.beauchamp@choderalab.org; john.chodera@choderalab.org

[*]To whom correspondence should be addressed

[†]Computational Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY

[‡]Tri-Institutional Program in Computational Biology and Medicine, Weill Cornell Medical College, New York, NY

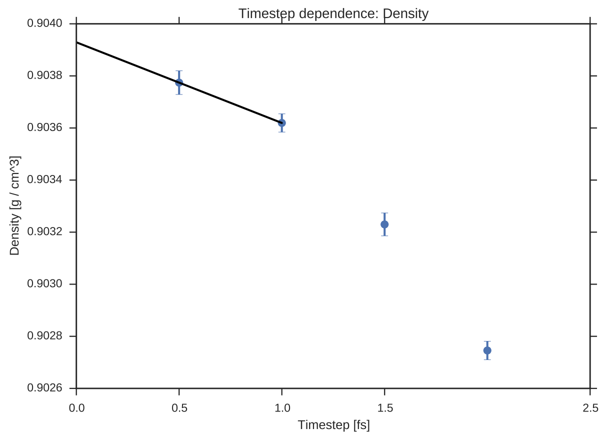
[¶]Graduate Program in Physiology, Biophysics, and Systems Biology, Weill Cornell Medical College, New York, NY

[§]OpenEye Scientific Software Inc., Santa Fe, NM

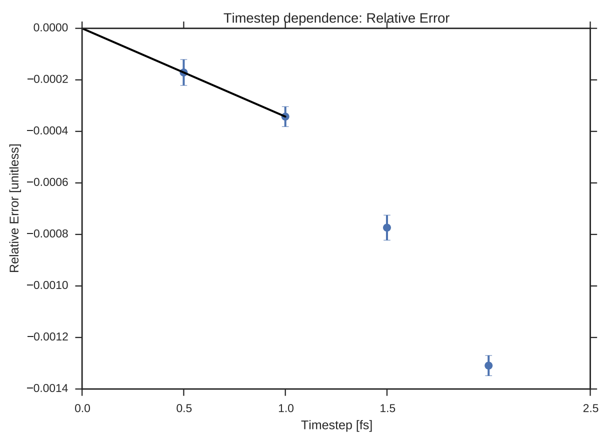
[‖]Thermodynamics Research Center, NIST, Boulder, CO

1 Supporting Information

- Figure: Timestep-dependence of density
- Figure: Error analysis (density) for ThermoML dataset
- Figure: Error analysis (static dielectric constant) for ThermoML dataset
- Figure: Temperature Dependence: Density
- Figure: Temperature Dependence: Static Dielectric Constant
- Commands to install dependencies

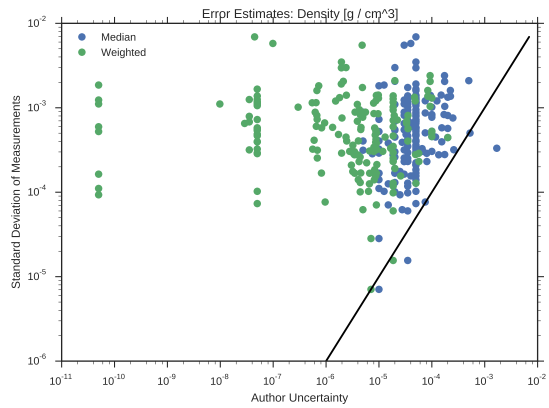


(a)

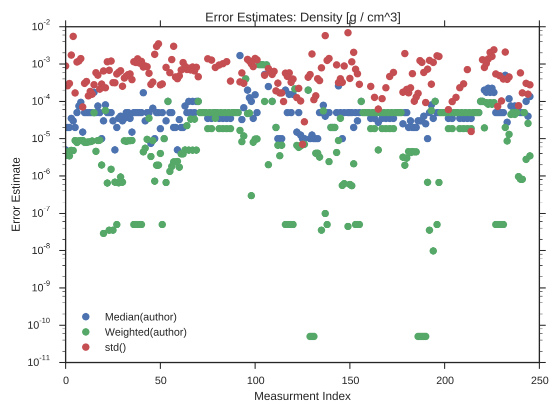


(b)

Figure S1: **Dependence of computed density on simulation timestep.** To probe the systematic error from finite time-step integration, we examined the timestep dependence of butyl acrylate density. (a). The density is shown for several choices of timestep. (b). The relative error, as compared to the reference value, is shown for several choices of timestep. Error bars represent standard errors of the mean, with the number of effective samples estimated using pymbar's statistical inefficiency routine.[1] The reference value is estimated by linear extrapolation to 0 fs using the 0.5 fs and 1.0 fs data points; the linear extrapolation is shown as black lines. We find a 2 fs timestep leads to systematic biases in the density on the order of 0.13%, while 1 fs reduces the systematic bias to approximately 0.03%—we therefore selected a 1 fs timestep for the present work, where we aimed to achieve three digits of accuracy in density predictions.

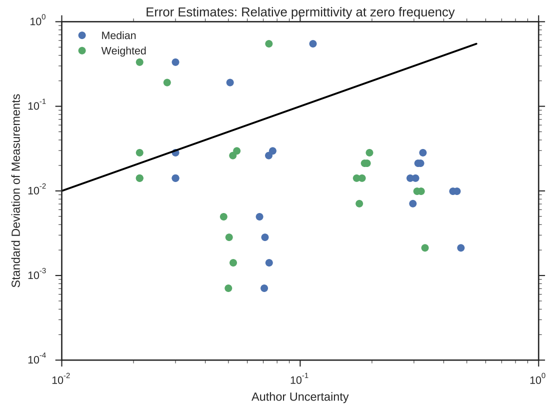


(a)

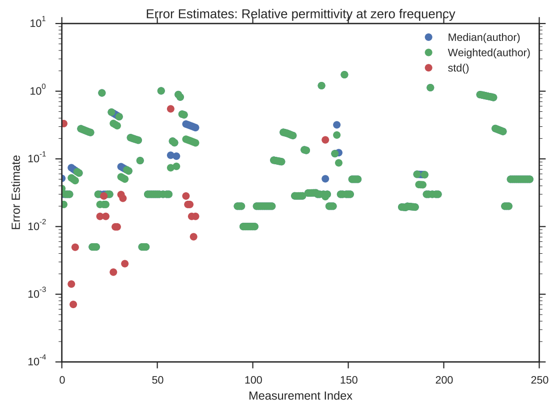


(b)

Figure S2: **Assessment of experimental error: Density** To assess the experimental error in our ThermoML extract, we compared three different estimates of uncertainty. In the first approach (Weighted), we computed the standard deviation of the optimally weighted average of the measurements, using the uncertainties reported by authors ($\sigma_{\text{Weighted}} = [\sum_k \sigma_k^{-2}]^{-0.5}$). This uncertainty estimator places the highest weights on measurements with small uncertainties and is therefore easily dominated by small outliers and uncertainty under-reporting. In the second approach (Median), we estimated the median of the uncertainties reported by authors; this statistic should be robust to small and large outliers of author-reported uncertainties. In the third approach (Std), we calculated at the standard deviation of independent measurements reported in the ThermoML extract, completely avoiding the author-reported uncertainties. Plot (a) compares the three uncertainty estimates. We see that author-reported uncertainties appear to be substantially smaller than the scatter between the observed measurements. A simple psychological explanation might be that because density measurements are more routine, the authors simply report the repeatability stated by the manufacturer (e.g., 0.0001 g/cm^3 for a Mettler Toledo DM40[2]). However, this hardware limit is not achieved due to inconsistencies in sample preparation and experimental conditions; see Appendix in Ref.[3] Panel (b) shows the same information as (a) but as a function of the measurement index, rather than as a scatter plot—because not all measurements have author-supplied uncertainties, panel (b) contains slightly more data points than (a).



(a)



(b)

Figure S3: **Assessment of experimental error: Static Dielectric Constant** To assess the experimental error in our ThermoML extract, we compared three different estimates of uncertainty. In the first approach (Weighted), we computed the standard deviation of the optimally weighted average of the measurements, using the uncertainties reported by authors ($\sigma_{\text{Weighted}} = [\sum_k \sigma_k^{-2}]^{-0.5}$). This uncertainty estimator places the highest weights on measurements with small uncertainties and is therefore easily dominated by small outliers and uncertainty under-reporting. In the second approach (Median), we estimated the median of the uncertainties reported by authors; this statistic should be robust to small and large outliers of author-reported uncertainties. In the third approach (Std), we calculated at the standard deviation of independent measurements reported in the ThermoML extract, completely avoiding the author-reported uncertainties. Plot (a) compares the three uncertainty estimates. Unlike the case of densities, author-reported uncertainties appear to be somewhat larger than the scatter between the observed measurements. Panel (b) shows the same information as (a) but as a function of the measurement index, rather than as a scatter plot—because not all measurements have author-supplied uncertainties, panel (b) contains slightly more data points than (a).

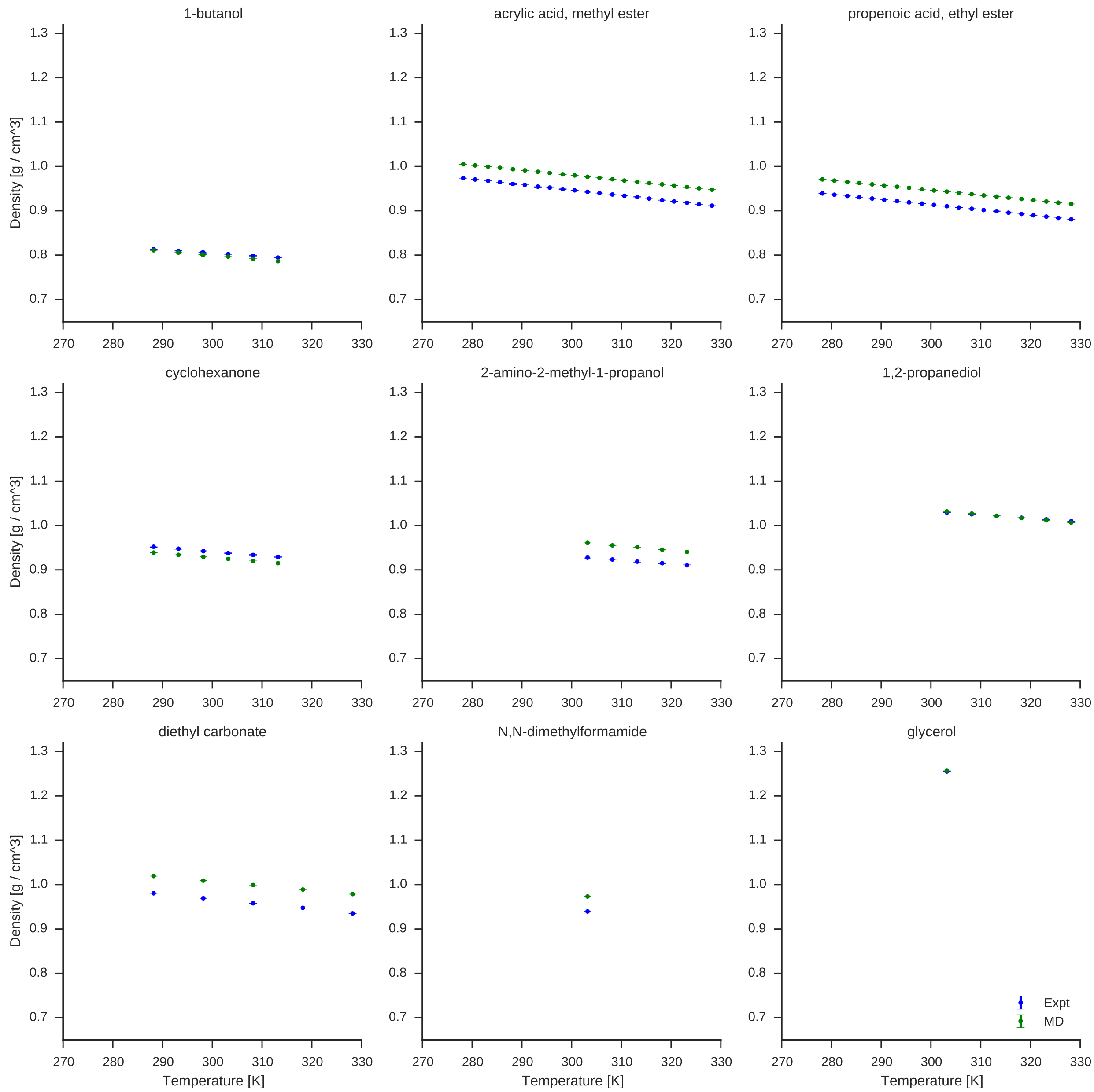


Figure S4: **Comparison of simulated and experimental densities for all compounds.** Measured (blue) and simulated (green) densities are shown in units of g/cm^3.

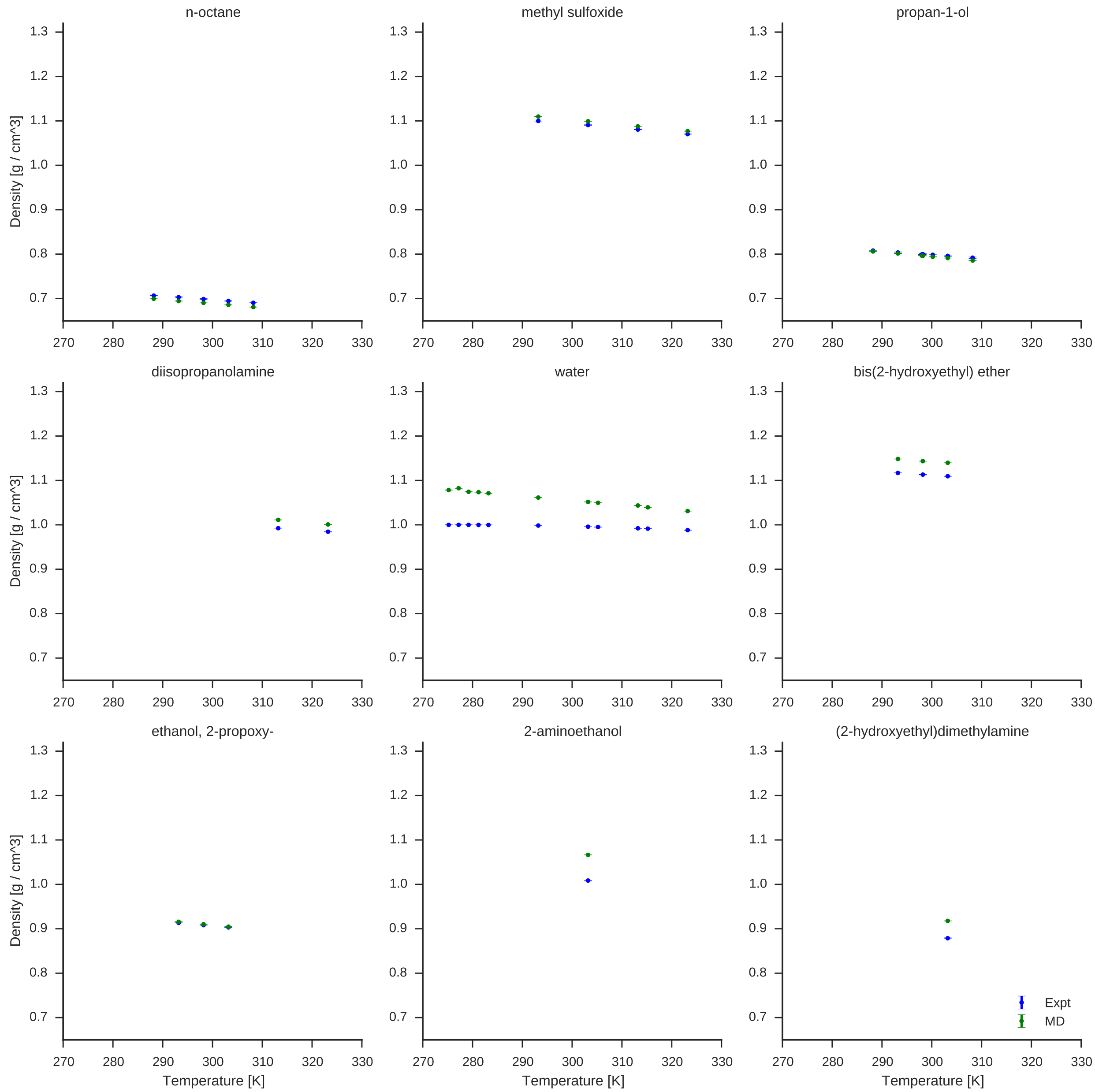


Figure S4: **Comparison of simulated and experimental densities for all compounds.** Measured (blue) and simulated (green) densities are shown in units of g/cm^3.

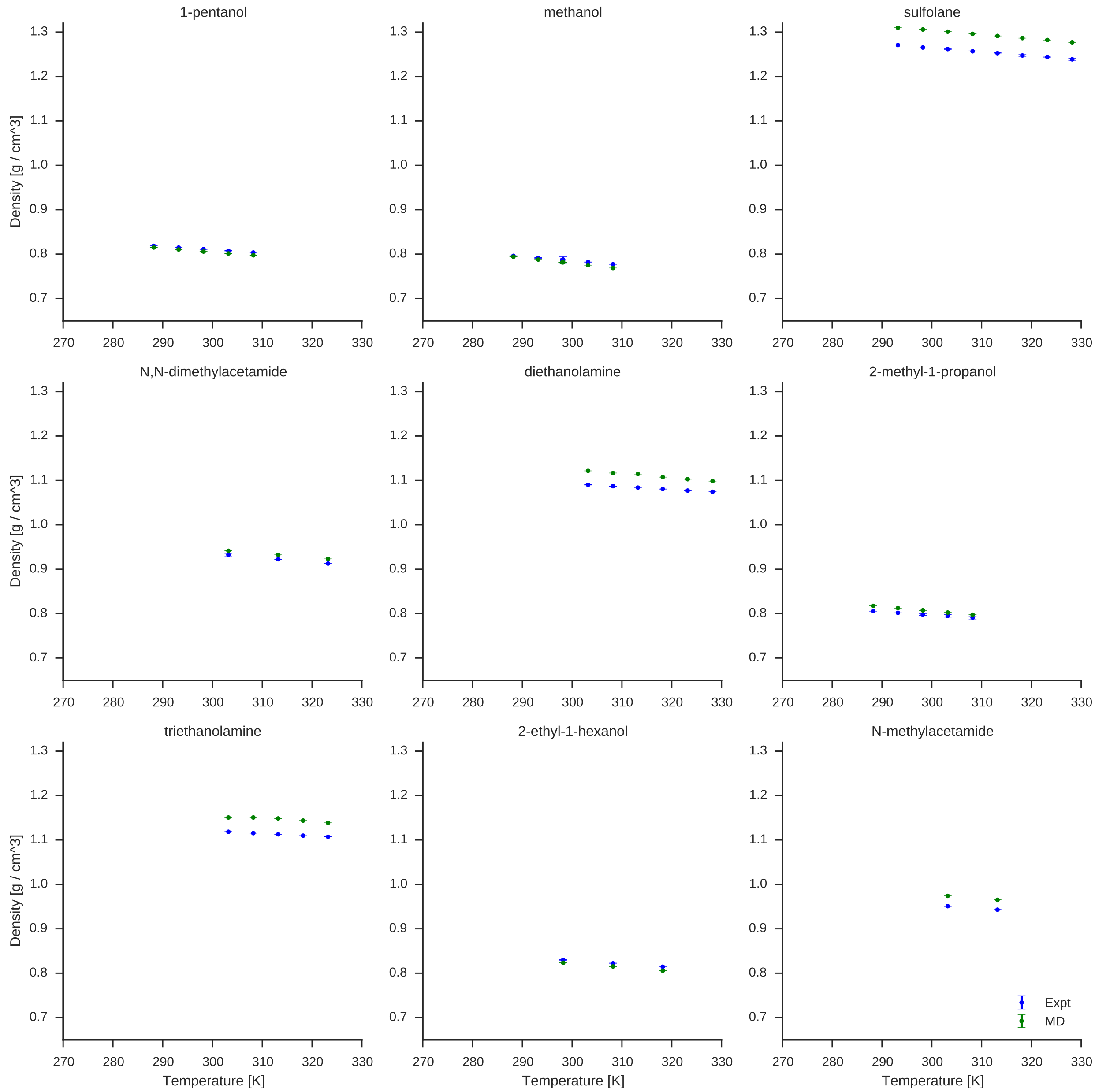


Figure S4: **Comparison of simulated and experimental densities for all compounds.** Measured (blue) and simulated (green) densities are shown in units of g/cm^3.

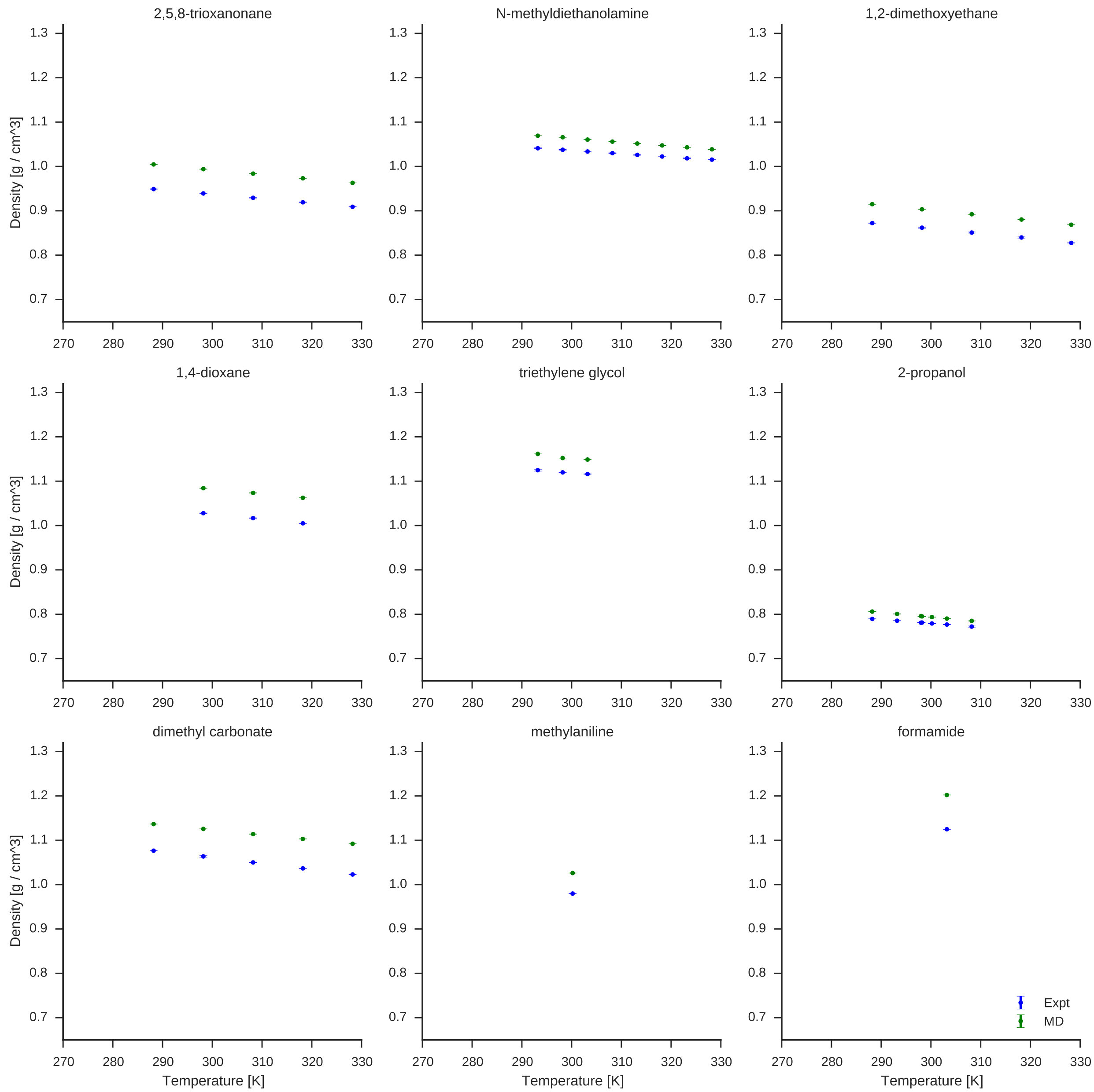


Figure S4: **Comparison of simulated and experimental densities for all compounds.** Measured (blue) and simulated (green) densities are shown in units of g/cm^3.

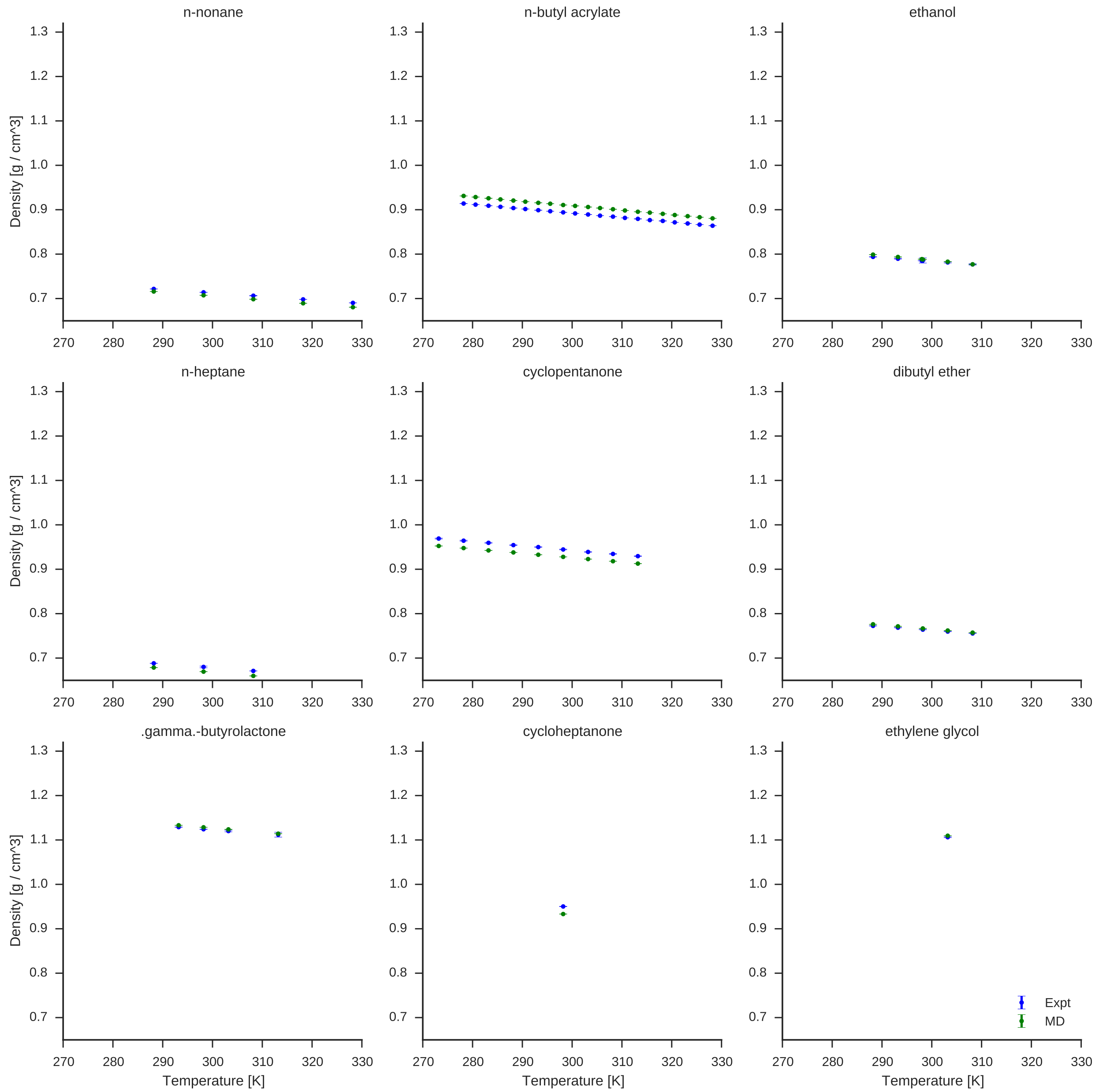


Figure S4: **Comparison of simulated and experimental densities for all compounds.** Measured (blue) and simulated (green) densities are shown in units of g/cm^3.

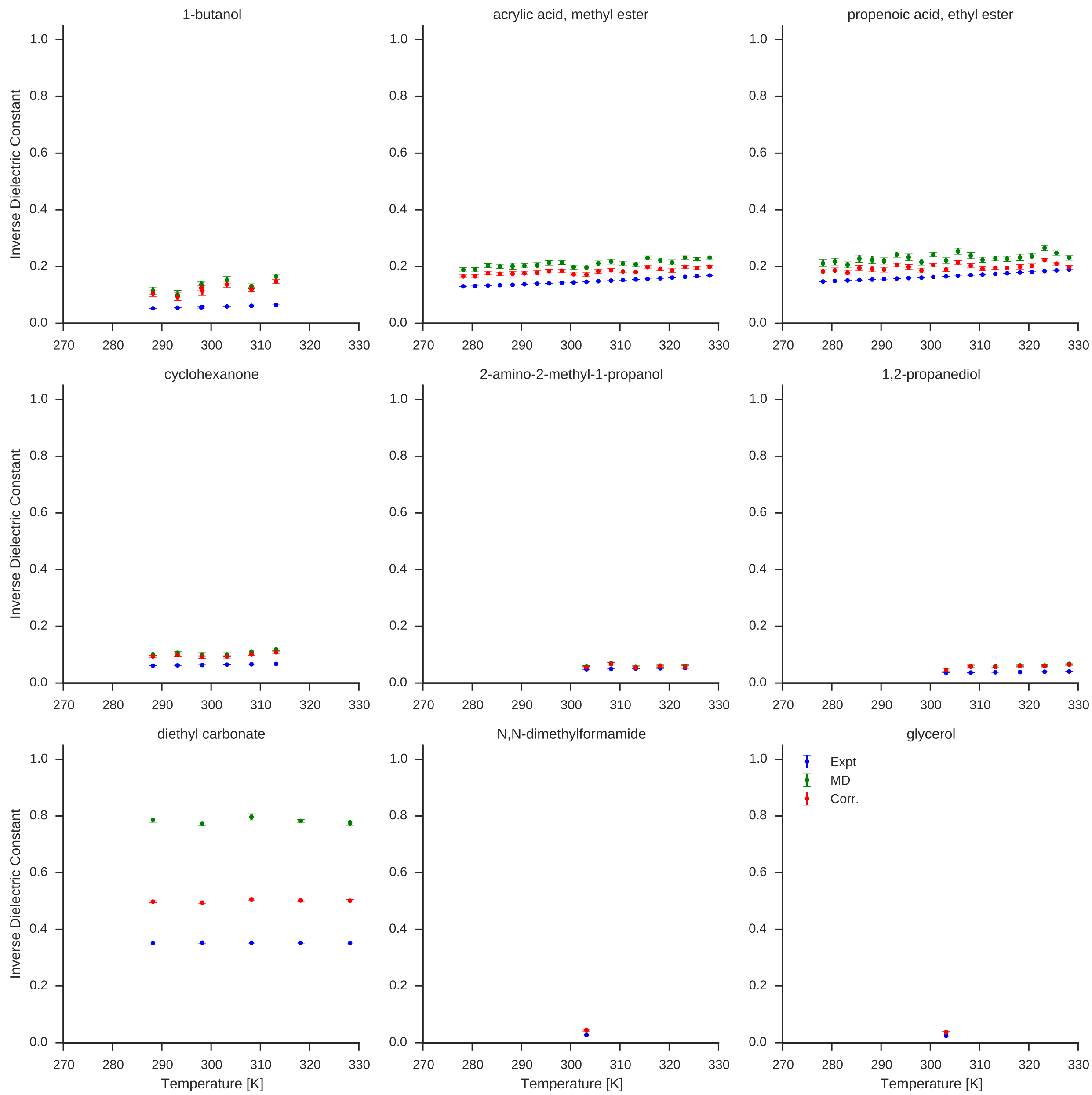


Figure S5: **Comparison of simulated and experimental inverse static dielectric constants for all compounds.** Measured (blue), simulated (green), and polarizability-corrected simulated (red) inverse static dielectric constants are shown for all compounds.

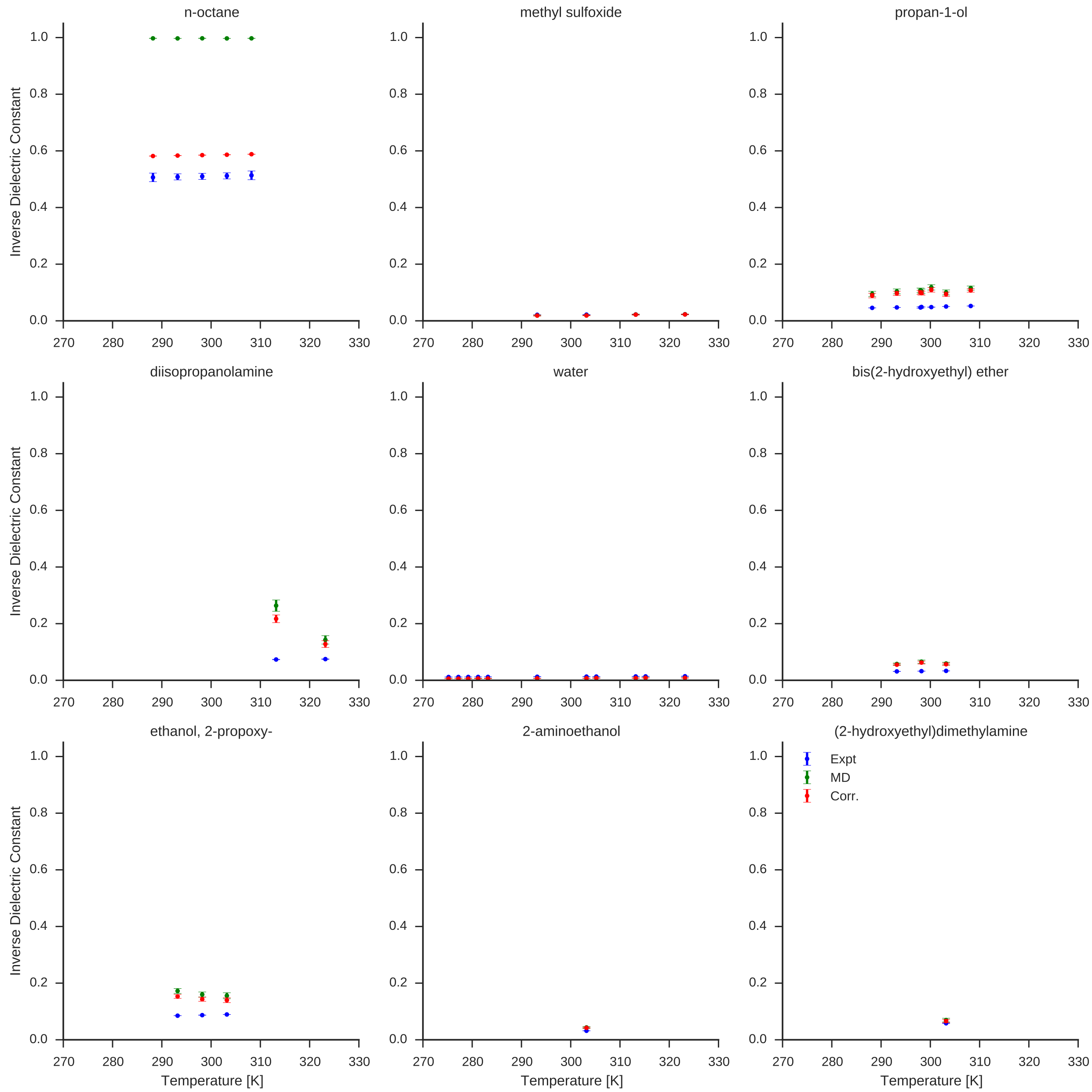


Figure S5: **Comparison of simulated and experimental inverse static dielectric constants for all compounds.** Measured (blue), simulated (green), and polarizability-corrected simulated (red) inverse static dielectric constants are shown for all compounds.

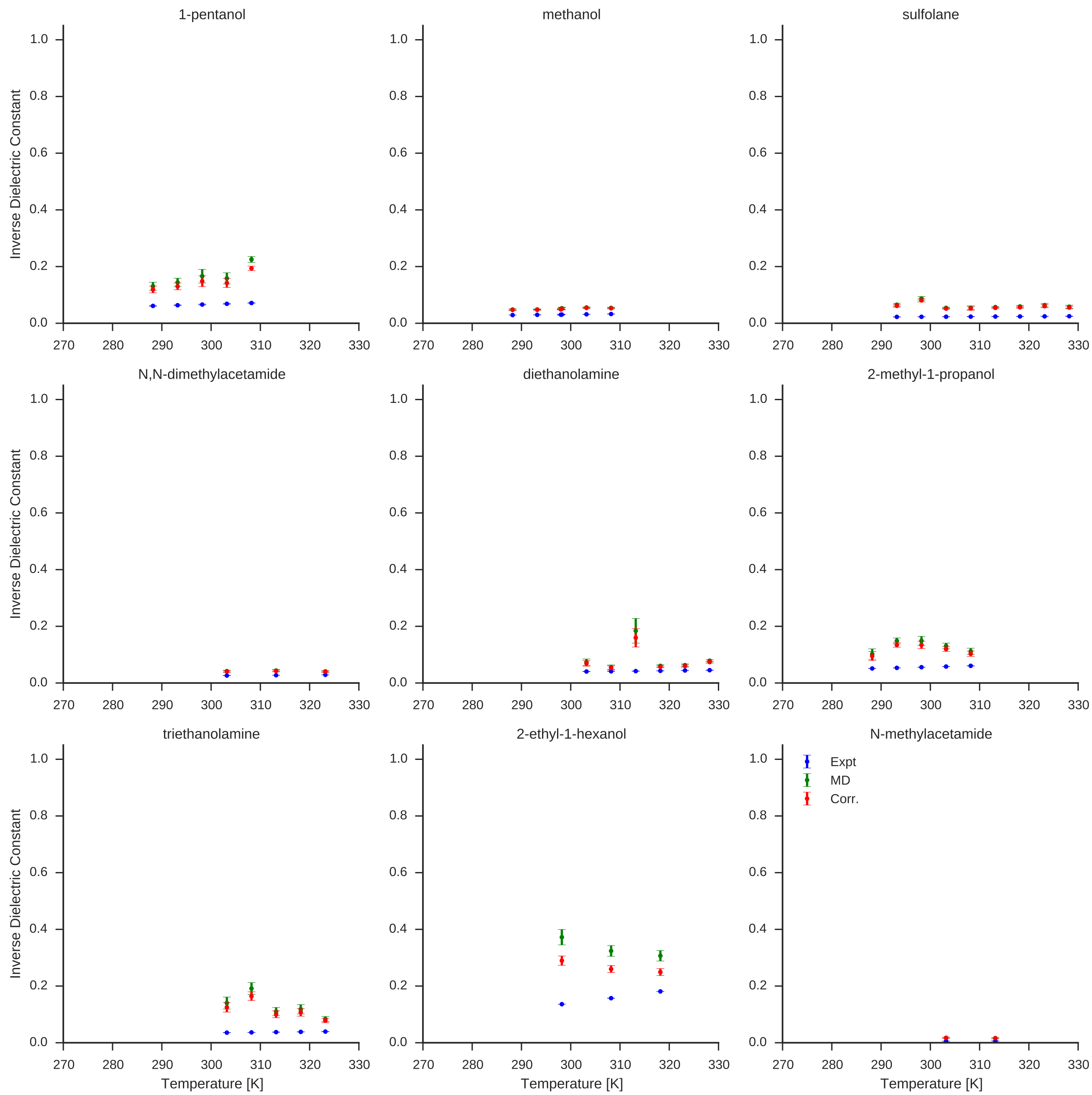


Figure S5: **Comparison of simulated and experimental inverse static dielectric constants for all compounds.** Measured (blue), simulated (green), and polarizability-corrected simulated (red) inverse static dielectric constants are shown for all compounds.

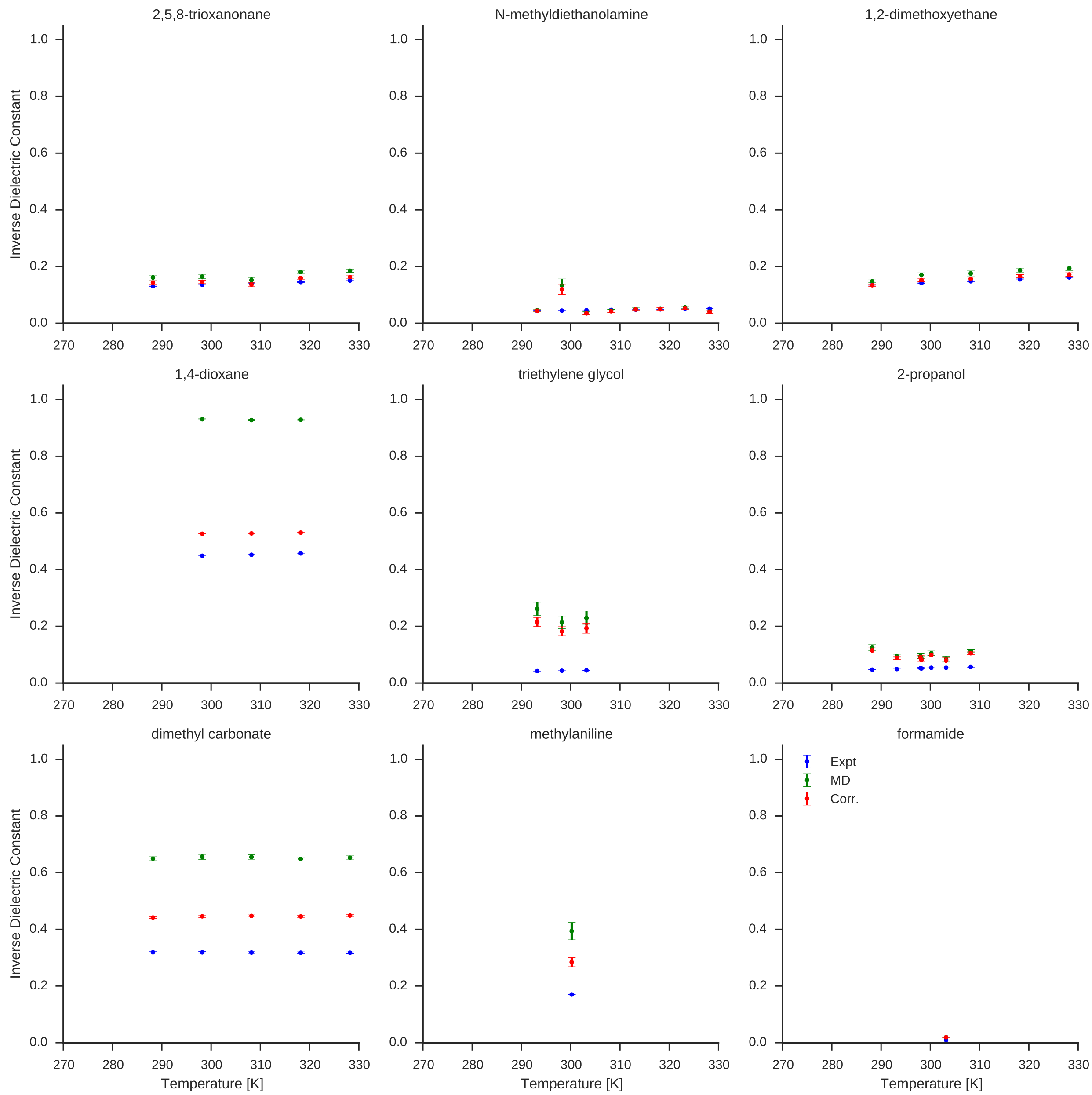


Figure S5: **Comparison of simulated and experimental inverse static dielectric constants for all compounds.** Measured (blue), simulated (green), and polarizability-corrected simulated (red) inverse static dielectric constants are shown for all compounds.

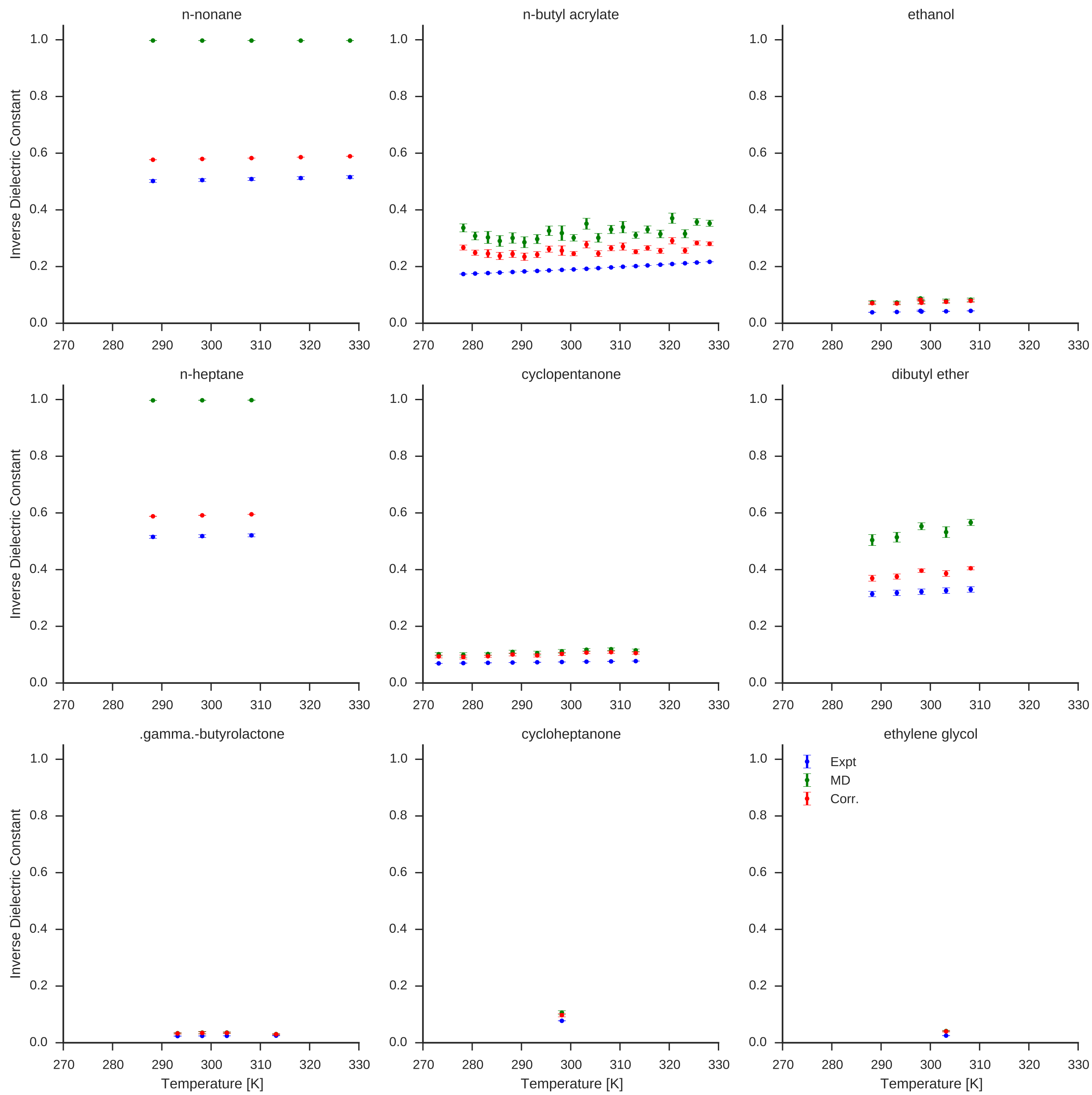


Figure S5: **Comparison of simulated and experimental inverse static dielectric constants for all compounds.** Measured (blue), simulated (green), and polarizability-corrected simulated (red) inverse static dielectric constants are shown for all compounds.

1.1 Dependency Installation

The following shell commands can be used to install the necessary prerequisites via the `conda` package manager for Python:

```
$ conda config −−add channels http://conda.binstar.org/omnia
$ conda install "openmoltools" "pymbar==2.1" "mdtraj==1.3" "openmm==6.3" packmol
%
```

Note that this command installs the exact versions used in the present study, with the exception of openmoltools for which only a more recent package is available. However, for authors interested in extending the present work, we suggust using the most up-to-date versions available instead, which involves replacing the equality symbols == with >=.

References

(1) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, 124105.

(2) Mettler Toledo Density Meters. `http://us.mt.com/us/en/home/products/Laboratory_Analytics_Browse/Density_Family_Browse_main/DE_Benchtop.tabs.models-and-specs.html`, [Online; accessed 15-Jan-2015].

(3) Chirico, R. D.; Frenkel, M.; Magee, J. W.; Diky, V.; Muzny, C. D.; Kazakov, A. F.; Kroenlein, K.; Abdulagatov, I.; Hardin, G. R.; Acree Jr, W. E. Improvement of Quality in Publication of Experimental Thermophysical Property Data: Challenges, Assessment Tools, Global Implementation, and Online Support. *J. Chem. Eng. Data* **2013**, *58*, 2699–2716.