

Assessing the quality of molecular simulations for vapor–liquid equilibria: An analysis of the TraPPE database

Becky L. Eggimann,^{*,†} Yangzesheng Sun,[‡] Robert F. DeJaco,^{‡,¶} Ramanish
Singh,^{‡,¶} Muhammad Ahsan,[‡] Tyler R. Josephson,[‡] and J. Ilja Siepmann^{*,‡,¶}

[†]*Department of Chemistry, Wheaton College, Wheaton, IL, 60187-5501, United States*

[‡]*Department of Chemistry and Chemical Theory Center,
University of Minnesota, Minneapolis, MN, 55455-0431, United States*

[¶]*Department of Chemical Engineering and Materials Science,
University of Minnesota, Minneapolis, MN, 55455-0431, United States*

E-mail: becky.eggimann@wheaton.edu; siepmann@umn.edu

Abstract

As molecular modeling and simulation techniques become increasingly important sources of thermophysical property and phase equilibrium data, the ability to assess the robustness of that data becomes more critical. Recently, the use of the compressibility factor (Z) has been suggested as a metric for testing the quality of simulation data for vapor–liquid equilibria (VLE). Here, we analyze predicted VLE data from the transferable potentials for phase equilibria (TraPPE) database and show that, apart from data entry or typographical errors, Z will always be well-behaved in Gibbs ensemble Monte Carlo (GEMC) simulations even when the simulations are not sufficiently equilibrated, but the same does not hold true for grand canonical Monte Carlo simulations.

When the pressure is calculated from the internal forces, then pressure and density are strongly correlated for the vapor phase and, for GEMC simulations, it is recommended to treat Z as an instantaneous mechanical property. From analysis of the TraPPE VLE data, we propose a complementary metric based on the predicted vapor pressures at three neighboring temperatures and their deviation from a local Clausius-Clapeyron fit.

Introduction

Molecular modeling and simulation (MMS) techniques, as tools for research and data collection, are maturing. Most obviously, computing power has increased dramatically since the first simulations were performed. Sophisticated algorithms and varying levels of abstraction in molecular models—from interactions described at the level of Kohn-Sham density functional theory to interactions of coarse-grained supra-atoms or even multi-molecule beads—have allowed MMS to reach an ever widening range of applications. MMS approaches are routinely used to calculate thermophysical properties and phase equilibrium data, and these data sources are becoming increasingly important as alternatives to experiment.¹ Furthermore, software packages are available that provide non-experts with relatively easy access to MMS methods. In a recent review, one of the pioneers of MMS noted that, where before researchers had to think carefully about how to perform their simulations because they were expensive, they now must think carefully because they have become cheap (mostly).² The barriers for the use of MMS are already low and continue to decrease. When this ease of use is combined with broad applicability, carelessness leading to low-quality data becomes a real danger.^{1–4}

There are a large number of factors that contribute to the quality of simulation data. These span the choices of models and algorithms, their implementation into a simulation program, the choice of experimental data for (and approach to) validation, the myriad of small details involved in correctly setting up the simulation algorithm, and the methods used

for analysis and averaging of properties. With so many critical decisions to make, expertise and experience will always be important for the proper execution of MMS techniques. As the use of MMS grows, it becomes more important than ever to complement experience with careful analysis of metrics, or “auxiliary quantities,”^{5,6} that can be monitored to assess simulation quality. A recent publication in this journal demonstrates such an approach for vapor–liquid equilibria (VLE). In that study, Nezbeda⁷ showcases an unfortunate lack of quality in a number of published VLE data, and suggests the use of the compressibility factor (Z) as a metric for uncovering these mistakes before the data are published. We want to affirm and support this approach, and others like it,^{2,3} but here offer some necessary additions and clarifications that are particularly relevant for VLE data generated using Gibbs ensemble Monte Carlo (GEMC) simulations.^{8,9}

Nezbeda’s own interests led him to calculate the compressibility factor for the vapor phase data from several published VLE studies using common water models.⁷ The results were sufficiently poor (i.e., the vapor data revealed significant outliers and weak agreement with experiment for compressibility factors along the vapor–liquid coexistence curve, see Figures 1–3 in the original work⁷) that further compounds were investigated, leading ultimately to a not insignificant number of poorly represented vapor phases.⁷ As Nezbeda indicates, the primary interest in VLE simulations is often the liquid phase and interfacial properties.⁷ In our own work, developing models for the transferable potentials for phase equilibria (TraPPE) force field by fitting to VLE data, we have historically allowed less relative precision/accuracy in the vapor-phase density compared to that of the liquid phase (less than 10% relative error for the vapor versus 1% for the liquid).¹⁰ Still, the frequency and type of errors noted by Nezbeda are perhaps unexpected and indicate more attention must be paid to the vapor phase in these important simulations. Interestingly, the first two TraPPE publications^{11,12} reported numerical values of the compressibility factor, but this practice was unfortunately not continued in subsequent TraPPE publications.

The data analyzed by Nezbeda span three common simulation techniques: Monte Carlo

or molecular dynamics simulations using an elongated simulation box with two explicit interfaces, GEMC simulations, and grand canonical Monte Carlo (GCMC) simulations.⁷ All compressibility factors were calculated from the data available for the saturated vapor pressure and the vapor densities. However, utility of Z as a metric will be affected by the differing methods used to determine pressure for each simulation technique. In particular, GEMC simulations employ the virial equation (Eqn. 1) to obtain the instantaneous pressure as a mechanical property for a given configuration of a given box (usually, data are reported only for the vapor phase) that is then averaged throughout the course of the simulation where the number of molecules, N_{box} , and the volume of the box, V_{box} , of interest fluctuate:

$$\begin{aligned}
 P_{\text{GEMC}} &= \left\langle \frac{N_{\text{box}} k_{\text{B}} T}{V_{\text{box}}} - \frac{1}{3V_{\text{box}}} \left(\sum_{i=1}^{N_{\text{box}}} \mathbf{r}_i \cdot \mathbf{f}_i \right) \right\rangle \\
 &= \left\langle \frac{N_{\text{box}} k_{\text{B}} T}{V_{\text{box}}} - \frac{1}{3V_{\text{box}}} \left(\sum_{i=1}^{N_{\text{box}}-1} \sum_{j=i+1}^{N_{\text{box}}} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} \right) \right\rangle
 \end{aligned} \tag{1}$$

The first term represents the contribution from the ideal gas law and the second term, the virial term, depends on the total internal force, $\mathbf{f}_i$, acting on particle i at position $\mathbf{r}_i$ or, for pairwise-additive potentials, on the intermolecular forces between atoms i and j in the box. As the number density decreases and the mean separation between molecules increases, the forces diminish and, consequently, the second term drops out and the instantaneous pressure becomes equal to that of an ideal gas at this density. Given this method of calculating the pressure, the compressibility factor, which is just the ratio of the molar volume of a real gas to that of an ideal gas, has to approach unity whenever the number density becomes small (assuming negligible aggregation). At low reduced temperatures ($T_{\text{r}} = T/T_{\text{crit}}$), the vapor phase of a GEMC simulation will have weak molecule–molecule interactions due to the low density and its pressure will be that of an ideal gas, i.e., P , calculated using Equation 1, must approach P_{ideal} , and Z must approach unity. This implies that when performed correctly, GEMC simulations should always produce well-behaved compressibility factors at low T_{r} (the same temperature region where Nezbeda saw the largest deviations in the data he

analyzed⁷). Of particular concern is that this behavior for Z is always expected, even when GEMC simulations under-sample the vapor phase by a phase ratio that leads to a very small average number of particles in the vapor, or have not reached equilibrium due to insufficient accepted transfer moves between the phases. Thus, Z is not a useful metric for detecting common sampling problems in GEMC simulations. In contrast, the pressure calculation for GCMC simulations with the histogram reweighting (HR) approach is based on the partition function, Ξ :

$$P_{\text{GCMC}} = \frac{k_{\text{B}}T}{V} \ln \Xi(\mu, V, T) \quad (2)$$

Thus, insufficient sampling in a GCMC simulation can lead to an erroneous pressure (and therefore an erroneous Z) even for a low-density vapor phase.

Here we calculate the compressibility factors for a large set of VLE data, that of the TraPPE database,¹³ and evaluate the use of Z as a metric for simulation quality. We also present a complementary metric based on the deviation of pressure triads (composed of three neighboring data points) from the Clausius-Clapeyron relationship, which is expected to do a better job of detecting sampling issues in GEMC simulations. For both metrics (Z -based and P -based), we develop quantitative measures of the extent of deviation from expected Z and P values and relate this to simulation quality; more deviation means less robust simulation data. The bounds of these quantitative measures are based on validation data¹⁴ obtained from longer simulations for larger system sizes for the seventeen compounds listed in Table 1.

Table 1: **Chemical Compounds, Models, and Purity^a**

chemical name	chemical or linear formula	CAS number	model
oxygen	H ₂ O	7782-44-7	TraPPE-small
carbon dioxide	CO ₂	124-38-9	TraPPE-small
ethane	C ₂ H ₆	74-84-0	TraPPE-UA
<i>n</i> -butane	CH ₃ (CH ₂) ₂ CH ₃	106-97-8	TraPPE-UA
<i>n</i> -decane	CH ₃ (CH ₂) ₈ CH ₃	124-18-5	TraPPE-UA
2-methylpropane	CH(CH ₃) ₃	75-28-5	TraPPE-UA
2,2-dimethylpropane	C(CH ₃) ₄	463-82-1	TraPPE-UA
2-methylpropene	(CH ₃) ₂ C=CH ₂	115-11-7	TraPPE-UA
2-methyl-1,3-butadiene	CH ₂ =CHC(CH ₃)=CH ₂	78-79-5	TraPPE-UA
naphthalene	C ₁₀ H ₈	91-20-3	TraPPE-UA
ethanol	CH ₃ CH ₂ OH	64-17-5	TraPPE-UA
2-butanol	CH ₃ CH(OH)CH ₂ CH ₃	78-92-2	TraPPE-UA
ethanal	CH ₃ CHO	75-07-0	TraPPE-UA
pentanal	CH ₃ (CH ₂) ₃ CHO	110-62-3	TraPPE-UA
acetone	CH ₃ COCH ₃	67-64-1	TraPPE-UA
2-methoxy-2-methylpropane	(CH ₃) ₃ COCH ₃	1634-04-4	TraPPE-UA
methyl acetate	CH ₃ COOCH ₃	79-20-9	TraPPE-UA
2-butanethiol	CH ₃ CH ₂ CH(SH)CH ₃	513-53-1	TraPPE-UA
dimethyl sulfide	(CH ₃) ₂ S	75-18-3	TraPPE-UA
<i>n</i> -perfluoropentane	CF ₃ (CF ₂) ₃ CF ₃	678-26-2	TraPPE-UA

^a All chemical samples are pure as specified by their respective input files.

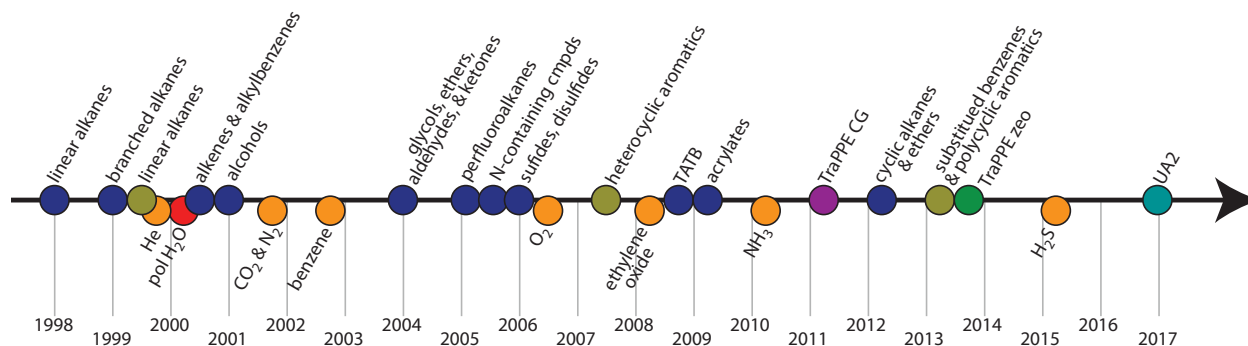


Figure 1: Timeline for the development of the TraPPE family of force fields. TraPPE families are represented as follows, in order from least to most coarse-graining: polarizable (red),³¹ all-atom small molecules (orange),^{26–30} explicit hydrogen (green),^{18,23–25} zeolites (bright green),³⁵ united atom-2nd generation “UA2” (teal),³⁴ united atom (blue),^{11,12,15–22} coarse grain (purple).³³

Computational Methods

The TraPPE Database

The development of the TraPPE family of force fields started with united-atom models for alkanes in 1998 and now includes (see Figure 1 for publication years and chemical functionalities): united-atom models for organic molecules containing C, H, N, and S atoms (TraPPE-UA),^{11,12,15–22} explicit-hydrogen models for alkanes, some aliphatic nitrogen-containing compounds, and aromatics (TraPPE-EH),^{18,23–25} all-atom models (including off-atom partial charges) for small molecules (TraPPE-small),^{26–30} polarizable models for water and methane (TraPPE-pol),^{31,32} coarse-grain models for hydrocarbons (TraPPE-CG),³³ second-generation united-atom models with off-atom sites (TraPPE-UA2),³⁴ and zeolites (TraPPE-zeo).³⁵ The TraPPE database,¹⁰ which can be accessed online through a dedicated website,¹³ is a collection of parameters and properties for several models and currently includes all TraPPE-UA models, some TraPPE-EH models, and all TraPPE-small models. One of the stated goals of the database and accompanying website is to provide users with supplemental information to facilitate the successful implementation and accurate use of TraPPE force fields. This necessarily includes force field parameters, simulation data, and additional documentation

Table 2: **Data Availability in the TraPPE Database**

Total Number of Molecules		Number of Data Points	Z	P Triads
with VLCC data	163	development data validation data	1072 209	754 169
with P_{vap} data	159			
with validation data	20			

of TraPPE’s conventions and unique features. A full description of the database and the available data has been previously published.¹⁰ Briefly, the available force field description consists of all functional forms and corresponding parameters for the bonded and non-bonded interactions of a given molecule, and properties are typically vapor–liquid coexistence densities, critical properties (T_{crit} , ρ_{crit} , and sometimes P_{crit}), saturated vapor pressures, and the normal boiling point. For our analysis here, we calculate the compressibility factor for all TraPPE models in the database using the previously published specific density and pressure of the vapor phase. We also use the vapor pressures to investigate the deviation of triads from the Clausius-Clapeyron relationship. Whereas the compressibility data are available at every temperature, the use of triads leads to $n_T - 2$ data points for a given compound (where n_T is the number of temperatures investigated for this compound). The resulting total numbers of data points for each metric are listed in Table 2.

Table 2 also references a validation dataset previously available only through the website¹⁴ but not yet published in the peer-reviewed literature. Data for these models (see Table 1) were determined from new simulations and confirm that previously developed parameters are still valid when used with much more rigorous simulation standards. For example, when ethane was originally parameterized in 1998, the simulation properties were calculated for a system of 400 molecules at six temperatures (from $T_r = 0.58$ to $T_r = 0.90$) by dividing a single independent run into 5 blocks, each 5000 cycles long (a cycle consists, in this case, of 400 Monte Carlo moves, or one move per particle in the simulation).¹¹ When ethane was simulated again in 2017,³⁴ the system size was increased to 1500 particles, eight temperatures were used (from $T_r = 0.58$ extending up to $T_r = 0.96$), the vapor phase was adjusted to make sure that 10-20% of the particles on average were in the vapor box for every temperature,

and rather than five blocks from a single run, system averages were calculated from eight truly independent runs, each 1,000,000 cycles in length. Numerical values of the validation data are provided in the Supporting Information.

Simulation Details

The work presented here includes 159 previously simulated TraPPE molecules and 20 newly simulated validation models. Complete simulation details for each previously published model can be found in the corresponding publications.^{11,12,15–25,31} In general, the majority of TraPPE models were simulated using coupled-decoupled configurational-bias Monte Carlo¹² in the Gibbs ensemble,^{8,9} though some models (especially in TraPPE-6¹⁷ and TraPPE-8¹⁹) used GCMC-HR techniques.^{36,37} As computational resources have become more accessible, system sizes have been increased, the simulations have been run longer, and more independent runs are utilized to compute average properties with statistical uncertainties. The TraPPE potential is similar to other molecular mechanics force fields and includes explicit terms for both bonded and nonbonded interactions of UA and EH beads. Bond lengths (1-2 interactions) are kept fixed in Monte Carlo simulations for all TraPPE models, bend angles (1-3 interactions) are governed by harmonic potentials, and torsions (1-4 interactions) use one of several functional forms as described in the original publications or on the website. Note that the website provides further detail on the specific conventions for defining the minimum-energy point for the dihedral angle, which varies depending on the functional form being used. For the models available in the TraPPE database, the nonbonded interactions are modeled with Lennard-Jones and Coulomb potentials. The Lennard-Jones potential is spherically truncated at 14 Å, with analytic tail corrections beyond the cutoff. When partial charges are included, the Ewald summation technique is employed for calculating the Coulomb interactions. Over time, the Ewald convergence parameter has also been increased to reflect gains in computing power.

All of the 20 validation models are simulated in the *NVT*-Gibbs ensemble with coupled-

decoupled configurational-bias Monte Carlo, with at least eight independent trajectories using the MCCC-S-MN software (Monte Carlo for Complex Chemical Systems-Minnesota).³⁸ System sizes are determined such that the liquid phase maintains a box length of more than 32 Å. At each temperature, the system volume is adjusted so that (on average) 10-20% of the molecules are in the vapor phase throughout the simulation. Each model is simulated at a minimum of eight temperatures that range from below the normal boiling point to very near the critical point. At least three, but usually four or more, temperatures are higher than $0.9T_{\text{crit}}$ and are used for estimating the critical properties. Simulations are also run until the relative standard errors of the mean are less than 0.5% for the liquid densities, 2% for the vapor pressures, 1% for the critical temperature, and 5% for the critical pressure. To achieve these standards, the simulations use optimized protocols for efficient GEMC simulations.^{39,40}

Results and Discussion

The TraPPE database represents one of the largest (if not the largest) available collections of VLE data determined by Monte Carlo simulations, and includes data from both the grand canonical and Gibbs ensembles. As Figure 1 shows, the available models cover a wide-range of chemical functionalities, span a number of years, and will naturally reflect the increasing simulation standards that have come with gains in computational power over time. As such, the TraPPE database contains an ideal collection of VLE data, especially useful for evaluating various metrics that might be used to judge simulation quality.

A Metric Based on the Compressibility Factor

The compressibility factor, Z , is a measure of the deviation shown by a real gas from ideal-gas behavior, and can be used to determine whether or not the vapor phase of a VLE simulation is well-behaved.^{7,41} As shown in Equation 3, Z is just the ratio of the molar volume of a real gas to that of an ideal gas, and can be calculated from the pressure and specific density of

the vapor phase in a VLE simulation:

$$Z = \frac{V_{\text{m,real}}}{V_{\text{m,ideal}}} = \frac{W/\rho}{RT/P} = \frac{P}{\rho} \frac{W}{RT} \quad (3)$$

Here, P is the pressure of the vapor in Pa, ρ is the specific density of the vapor in kg/m^3 , W is the molecular weight of the molecule in kg/mol , R is the gas constant in $\text{J}/(\text{K mol})$, and T is the absolute temperature in K.

For the vapor–liquid coexistence region, the expected behavior is that Z should be unity for low T_r (where both the density and pressure of the vapor are sufficiently low for the vapor to behave ideally) and decrease with increasing rapidity (as molecular interactions become increasingly important) toward some value of the critical compressibility factor. Robust VLE simulations for non-associating molecules should capture this trend, despite several published examples failing to do so, especially for low T_r .⁷ According to this metric, the more deviation from the expected trend in Z , the more likely it is that the simulations are flawed in some manner, thus yielding untrustworthy results. The amount of deviation from the expected compressibility is then a measure of simulation quality.

However, it should also be noted that for a few molecules, such as low-molecular-weight carboxylic acids and hydrogen fluoride, Z will deviate strongly from those of non-associating fluids at these conditions. There is a significant degree of association in the vapor-phase, even at low- T_r , due to the formation of hydrogen-bonded aggregates for these molecules. Their behavior differs from that of water and simple alcohols because their vapor pressures are relatively high due to their relatively low heats of vaporization resulting from similar numbers of hydrogen bonds in the liquid and vapor phases. Near the critical point, hydrogen-bond formation for water and alcohols also leads to a lowering of Z but to a lesser degree.

Even with these natural variations, the use of Z as a metric is especially appealing because it is easy to calculate and unexpected deviations should be immediately obvious. However, in simulations, the average compressibility can be calculated in two different ways and we

show here that these different approaches result in Z values that are systematically different and can lead to dramatic overestimations of the statistical uncertainty. Relying on Z as a metric requires further clarity about which approach was used to determine Z .

In the first approach, Z is calculated from ensemble-average values of P and ρ as utilized by Nezbeda.⁷ Statistical uncertainties would be calculated using error propagation from the resulting uncertainties in $\langle P \rangle$ and $\langle \rho \rangle$. This is called here the ensemble average approach (Z_e).

$$Z_e = \frac{\langle P \rangle}{\langle \rho \rangle} \frac{W}{RT} \quad (4)$$

The second approach calculates Z as a mechanical observable at fixed intervals throughout the simulation and determines errors directly from $\langle P/\rho \rangle$ for multiple independent runs (or block averages). This is the instantaneous average approach (Z_i).

$$Z_i = \left\langle \frac{P}{\rho} \right\rangle \frac{W}{RT} \quad (5)$$

To discuss the differences between these approaches, we denote the compressibility factor at each pressure calculation step as $z = P/(\rho RT)$ and evaluate the difference between Z_e and Z_i as

$$\begin{aligned} Z_e - Z_i &= \frac{\langle P \rangle}{\langle \rho \rangle RT} - \langle z \rangle \\ &= \frac{\langle z \rho RT \rangle}{\langle \rho \rangle RT} - \langle z \rangle \\ &= \frac{\langle z \rho \rangle}{\langle \rho \rangle} - \langle z \rangle \\ &= \frac{\langle z \rho \rangle - \langle z \rangle \langle \rho \rangle}{\langle \rho \rangle} \\ &= \frac{\text{cov}(z, \rho)}{\langle \rho \rangle} \end{aligned} \quad (6)$$

where $\text{cov}(\cdot, \cdot)$ is the covariance between two variables. The correlation between the instantaneous compressibility factor and density is directly proportional to the difference between

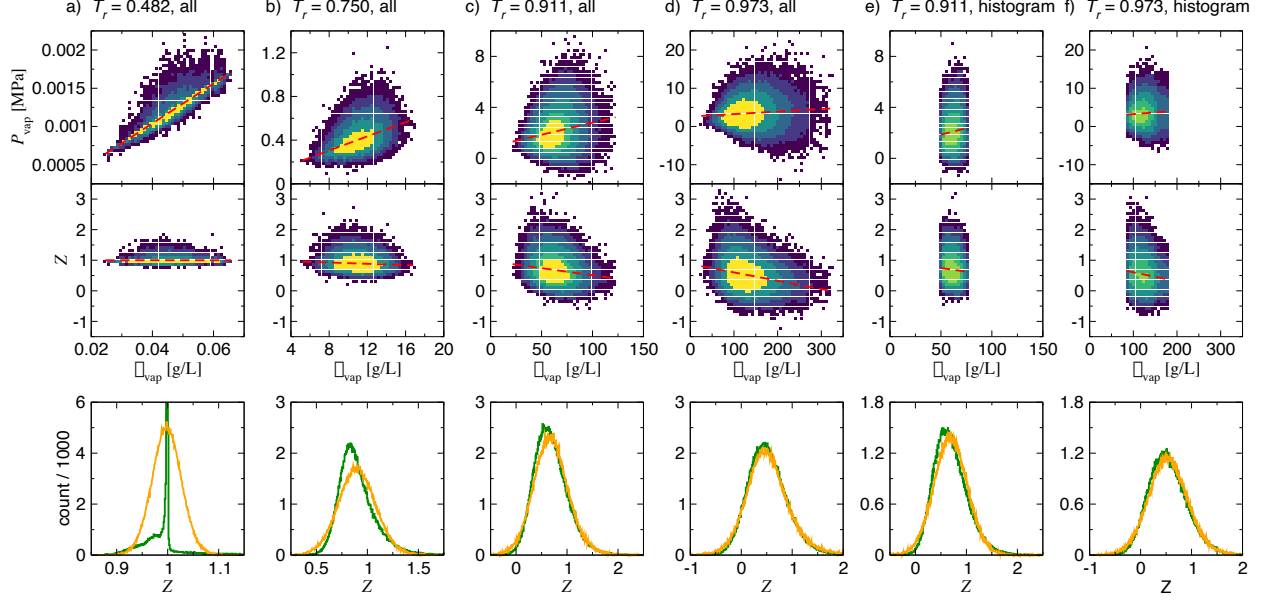


Figure 2: Distributions of P , ρ and z obtained from validation simulations for pentanal at T_r of (a) 0.482, (b) 0.750, (c) 0.911, and (d) 0.973, and columns (e) and (f) show data obtained after histogram analysis of the vapor density for $T_r > 0.9$. The top and middle rows show 2D distributions of (P, ρ) and (z, ρ) , respectively, with the dashed red lines showing the linear regressions. Color gradients from dark purple to yellow corresponds to the count of simulation frames N binned by $1 \leq N \leq 9$, $10 \leq N \leq 49$, $50 \leq N \leq 249$, $250 \leq N \leq 999$, $1000 < N \leq 2499$, and $N \geq 2500$. The bottom row shows histograms of compressibility factors from the simulation results (green) and from Gaussian maximum likelihood estimates that yield the same mean and covariance matrix as the simulation results (orange).

Z_e and Z_i .

To illustrate the differences between Z_e and Z_i , we use validation data for pentanal as an example (see Figure 2 and Table 3). Large fluctuations near T_{crit} may lead to interchange between liquid and vapor box identities, so larger system sizes (800 instead of 400 molecules) were used closed to T_{crit} to prevent switching and provide a robust comparison between Z_e and Z_i . At low T_r , the behavior of the vapor pressure is dominated by the ideal-gas contribution (see Eqn. 1) and, as a consequence, the compressibility factor is unity for a great majority of simulation steps. As a consequence, the vapor pressure is strongly correlated with the vapor density; the instantaneous pressure values are always positive and, at a given vapor density, the spread of the observed pressure values is smaller than a factor of two. However, the compressibility factor ($z \approx 1$) is nearly independent of the vapor density. Thus, at low

Table 3: **Discrepancies between Compressibility Factor Measurements Observed in Validation Simulations for Pentanal and their Gaussian Maximum Likelihood Estimates^a**

T_r	data	Z_e	Z_i	$R(P, \rho)$	$R(z, \rho)$	$(Z_e - Z_i)/Z_i$ simulation [%]	$(Z_e - Z_i)/Z_i$ Gaussian [%]
0.482	full	0.9981 ± 0.0004	0.9981 ± 0.0003	0.958 ± 0.015	-0.007 ± 0.014	-0.002 ± 0.005	-0.003 ± 0.006
0.750	full	0.898 ± 0.002	0.900 ± 0.002	0.479 ± 0.017	-0.072 ± 0.008	-0.18 ± 0.02	-0.19 ± 0.04
0.911	full	0.676 ± 0.007	0.685 ± 0.006	0.205 ± 0.007	-0.156 ± 0.014	-1.33 ± 0.19	-1.50 ± 0.28
0.973	full	0.502 ± 0.022	0.525 ± 0.021	0.084 ± 0.013	-0.232 ± 0.023	-4.3 ± 1.0	-6.0 ± 2.3
0.911	hist	0.692 ± 0.009	0.693 ± 0.009	0.104 ± 0.016	-0.072 ± 0.008	-0.28 ± 0.05	-0.27 ± 0.05
0.973	hist	0.541 ± 0.029	0.546 ± 0.029	0.052 ± 0.018	-0.103 ± 0.026	-0.82 ± 0.41	-0.86 ± 0.46

^a $R(\cdot, \cdot)$ denotes the Pearson correlation coefficient. All values were calculated from each of the 16 independent simulations and reported as the mean and standard deviation.

T_r , $\text{cov}(z, \rho)$ is small, and we observe $Z_e - Z_i \approx 0$. The Pearson correlation coefficient for P and ρ is greater than 0.96, whereas that for z and ρ is negative with a magnitude less than 0.02.

In contrast, at temperatures approaching the critical point, the vapor pressure is greatly influenced by mostly attractive, but also repulsive, interactions between molecules and, at a given vapor density, the pressure varies widely from positive to negative values. With interactions playing an important role, the vapor pressure is only very weakly but positively correlated with the vapor density ($R(P, \rho) \approx 0.05$ at $T_r = 0.973$). In such cases, the compressibility factor is found to exhibit a moderate negative correlation with vapor density (i.e., higher densities are more likely to encounter attractive interactions, see Figure 2), so that $Z_e < Z_i$.

Since vapor density and pressure are directly sampled from microscopic properties throughout the simulation trajectory, the instantaneous values of the compressibility factor follow a ratio distribution⁴² between the distributed density and pressure values. However, it is not feasible to analytically calculate the distribution of z given simple parametric models of ρ and P , such as a bivariate Gaussian distribution. Instead, Gaussian sampling experiments were conducted to gain insights into the distributions of simulation data and their influence on calculating the compressibility factor. Figure 2 shows histograms of z obtained from simulations versus estimations from Gaussian distributions matching the mean, variance, and

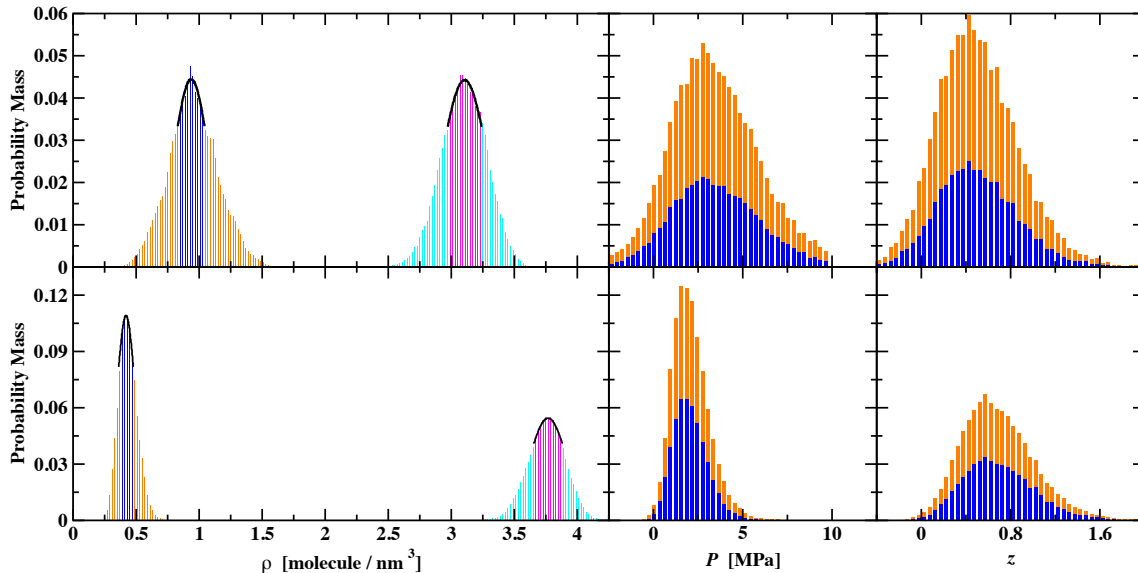


Figure 3: Probability mass histograms of the coexistence densities (left), vapor-phase pressures (middle), and compressibilities (right) obtained from the validation simulations for pentanal at $T_r = 0.911$ (bottom row) and 0.973 (top row). Gaussian distributions (black lines) are fit to blue and magenta bars with heights of at least 0.75 of the maximum height for the vapor and liquid coexistence densities, respectively. For pressures and compressibility factors, the blue bars illustrate the distributions of the instantaneous density points selected by the Gaussian fit approach for the vapor density, whereas the orange bars show data for the entire trajectories.

covariance of the simulation distributions. With increasing T_r , pressure and vapor density become less correlated, giving a larger difference between Z_e and Z_i . Although the z distributions are less similar to a Gaussian distribution at lower T_r , using a Gaussian distribution to model the fluctuations of P and ρ is still sufficient to qualitatively estimate the discrepancy between the two types of measurements for the compressibility factor (see Table 3). The Gaussian sampling experiments also indicate that, as T_{crit} is approached, the distribution of z broadens and Z_e gives a significantly smaller value than Z_i . The Gaussian sampling results further reveal the statistical prevalence of such discrepancies in the simulations, indicating that our analysis is also applicable to a wider range of compounds. Calculating Z as a mechanical property on-the-fly throughout GEMC simulations will avoid the error introduced through correlation of P and ρ .

For GEMC simulations close to the critical point, the distributions of the vapor and

liquid densities exhibit non-Gaussian behavior with tails toward intermediate densities.^{40,43} This tailing is a consequence of the coupling of the two phases and is more pronounced for smaller system sizes. Thus, to reduce finite-size effects, the distributions near the peaks for the vapor and liquid densities can be approximated by Gaussian distributions (only bins with a height of at least 0.75 of the respective maximum peak height are included for the Gaussian fit as recommended by Dinpajoo et al.⁴⁰). The mean values obtained from these Gaussian distributions can be taken as the best estimates of the vapor and liquid coexistence densities.⁴⁰ When the vapor pressure and compressibility factor for temperatures near the critical point are calculated, then only data from configurations within the regions used for the Gaussian fits to the density distributions should be considered. At $T_r > 0.9$, about 50% of the configurations meet the peak threshold. A Python script that performs this histogram analysis is described in the Supporting Information. Figure 3 shows probability mass histograms for the distributions of vapor and liquid densities, vapor-phase pressures, and compressibilities obtained from validation simulations for pentanal at $T_r = 0.911$ and 0.973. Due to the asymmetry of the density distribution, the histogram analysis removes a slightly larger contribution from the high-density tail than from the low-density tail of the ρ_{vap} distribution. These higher-density configurations are more likely to include significant attractive interactions and a lower P_{vap} and z (see Figure 2). Thus, the Z_i and Z_e values estimated from the histogram-selected configurations are slightly higher than those from the full trajectory (see Table 3). In addition, the histogram-selected data show smaller (in magnitude) $R(P, \rho)$ and $R(Z, \rho)$ values than the full trajectories and their $(Z_e - Z_i)/Z_i$ differences are described better by the Gaussian sampling experiments. To reduce finite-size effects near the critical point, our recommendation is to use the histogram analysis for the calculation of $\langle Z_i \rangle$.

Analysis of Compressibility Factors from the TraPPE Database

For our analysis of Z values in the TraPPE database, we, like Nezbeda, are relying on previously simulated data and must therefore calculate Z_e . When trajectories are available for some of the most recent validation data, then we utilize the preferred Z_i values. Figure 4 shows all of the the compressibility factors, calculated as Z_e , from the original development of UA, EH and small molecules in the development models available in the TraPPE database. We can immediately see outliers in some plots and unexpectedly large error bars in others, but before considering these deviations in greater detail, we first highlight a positive trend that can be easily overlooked. The individual plots on the left side of Figure 4 represent the passage of time, from the first TraPPE publication in panel A (linear alkanes,¹¹ 1998) to one of the more recent publications in panel I (cyclic alkanes and ethers,²² 2012). Over time, as the standards that can be achieved for VLE simulations have increased, the TraPPE data show a corresponding narrowing and smoothing of the Z curve, with decreasing error bars noted for individual data points. The TraPPE simulation data is improving in both accuracy (more consistent curves with fewer outliers) and precision (smaller error bars) as time progresses. This trend, which assures us that trustworthy data for VLE simulations can be achieved, stands in contrast to the examples highlighted by Nezbeda, and is worth emphasizing. While it is certainly becoming easier and easier to generate low-quality data given the rapid growth of computational methods and increasing access to simulation software for non-experts, the TraPPE compressibility factors demonstrate that this need not be the case. With appropriate expertise and careful attention to the (many) details of each simulation, excellent results can be achieved. For any metric to be useful it must be able to demonstrate both the praise-worthy and the cringe-worthy in our data, and inspire all simulators to strive for increasingly more robust simulation quality.

Returning, then, to a discussion of the outliers in Figure 4, we have identified four different types of errors that emerge when applying this metric—simple typos and data entry errors (examples highlighted with magenta arrows), molecular weight mistakes (the blue arrow next

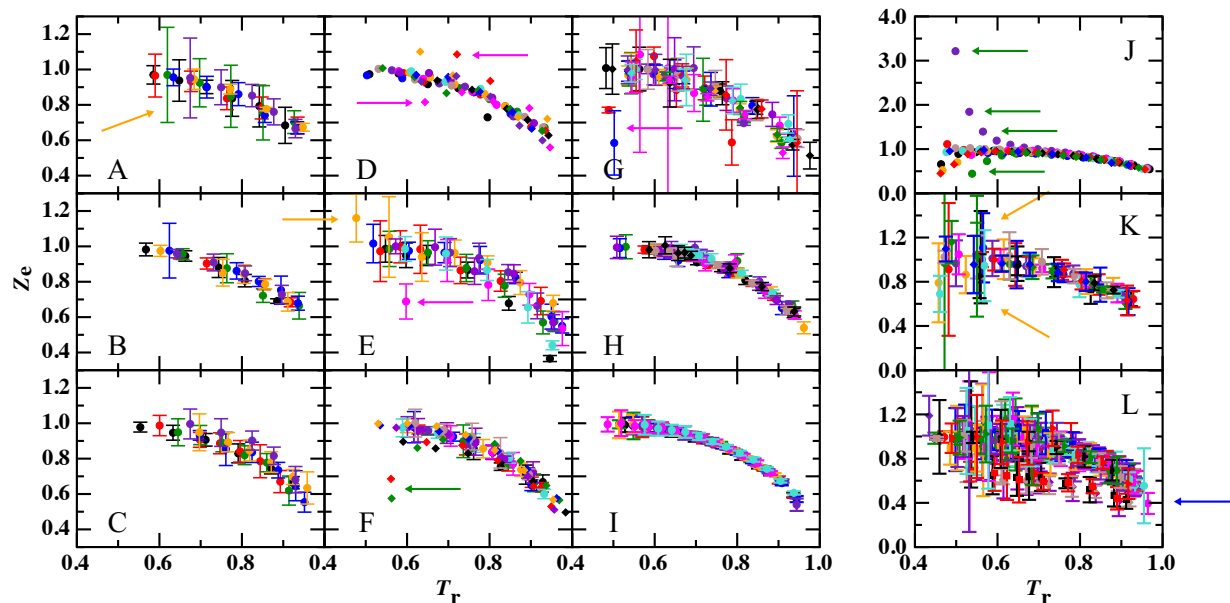


Figure 4: Compressibility factors from the TraPPE database. (A) linear alkanes, UA,¹¹ (B) branched alkanes, UA,¹² (C) linear alkanes, EH,²³ (D) alkenes and alkylbenzenes, UA,¹⁵ (E) alcohols, UA,¹⁶ (F) ethers, glycols, ketones, aldehydes, UA,¹⁷ (G) N-containing compounds, UA and EH,¹⁸ (H) acrylates, UA,²¹ (I) cyclic alkanes and ethers, UA,²² (J) sulfides, disulfides, thiophene, UA,¹⁹ (K) heterocyclic aromatics, EH,²⁴ (L) substituted benzenes, polycyclic aromatics, EH.²⁵ Different colored symbols represent different compounds. Colored arrows point to identified outliers, as described in the text.

to Panel L points to a set of data points that are systematically offset from the expected Z trend), insufficient sampling for GCMC-HR (green arrows in Panels F and J indicate data points that are rapidly curving away from the expected trend, either upwards towards higher and higher Z or downwards toward 0), and potential poor sampling in GEMC (orange arrows point to several examples of large error bars). The three panels offset to the right of Figure 4 contain outliers that require a larger scale than the other TraPPE data; panels K and L, representing data for EH models of aromatic compounds, also show the majority of the simulations with large error bars.

Typos and data entry errors show up as single points that deviate significantly from the observed trend for the rest of the data set. We were able to identify and correct a small number of data entry mistakes that appear in TraPPE publications as a result of this analysis. Corrected data are available at the TraPPE website.¹³ Some suspected errors—

those indicated with magenta arrows—could not be corroborated due to the inability to locate the original data (i.e., data collected before the era of data management plans).

Molecular weight mistakes lead to Z values that have the expected curve shape but are offset by a constant factor (even though the saturated vapor pressure may be correct; the vapor density is usually reported as specific density and not as number density or molar volume, see Eqn. 3). There are four molecules from TraPPE-10 with known molecular weight errors (Panel L, blue arrow)—anthracene, phenanthrene, naphthalene-2-carbonitrile, and naphthalene-2-ol. They have Z values that approach 0.7 (rather than unity) at low T_r . The TraPPE-EH models for aromatics use compound-specific partial charges on all ring atoms and substituent beads directly bonded to the ring, which are obtained from electronic structure calculations. These require separate entries in the parameter tables for our software and, throughout this work, we found numerous entries where the mass of a hydrogen atom was set to 0.012 kg/mol. These mistakes in correctly specifying the atomic weights also lead to specific liquid densities that are too high by a factor of about 1.4 ($= Z^{-1}$ at low T_r).

GCMC-HR simulations have characteristic errors that result from the particularities of that method. At low T_r , equilibration is especially challenging and the histogram-reweighting technique compounds any errors by including mistakes from insufficiently sampled low- T_r simulations in the histograms for the next higher temperatures, thereby creating the curved tails observed for these simulations (Panels F and J). Likely, it is only the simulation at the lowest T_r that is problematic, and some of the errors are indeed very large. GCMC-HR related deviations account for 20% of the outlying molecules (21% of the outlying data points), which is slightly larger than the total proportion of GCMC-HR simulations represented in the database (11% of molecules; 19% of data points).

Large error bars, such as in Figure 4, panels A, B, E, G, K, and L, are an indication of potential sampling problems in GEMC simulations. However, caution should be used when interpreting error bars for Z_e . The two values needed to calculate Z_e are P and ρ (see Eqn. 4), but we have already seen that these measures can be highly correlated in the

Table 4: **Error Estimates in Z_e and Z_i** ^a

Molecule	T_r	Z_e	Z_i	Error Ratio
pentanethiol	0.4983	0.99 ± 0.09	0.9970 ± 0.0004	230
	0.5316	1.00 ± 0.07	0.9945 ± 0.0005	140
	0.5648	0.99 ± 0.06	0.9890 ± 0.0010	60
	0.5980	0.99 ± 0.03	0.9830 ± 0.0010	30
dimethyl disulfide	0.4620	1.01 ± 0.16	0.9986 ± 0.0005	320
	0.4950	0.99 ± 0.09	0.9968 ± 0.0007	130
2-butanethiol	0.4668	1.01 ± 0.13	0.9983 ± 0.0006	220
	0.5027	0.99 ± 0.09	0.9961 ± 0.0005	180

^a Statistical uncertainties are reported as the standard error of the mean.

vapor phase at low T_r (see Figure 2). When using propagation of errors, applied at the end of a simulation, this correlation will lead to significant overestimations of the error in Z_e , as much as two orders of magnitude. To illustrate this point, new GEMC simulations were carried out at $T_r < 0.6$ for some sulfur-containing compounds. Table 4 shows examples of Z_e and Z_i values and their respective statistical uncertainties obtained from the same sets of independent simulations. Treating Z as a mechanical observable (Z_i) leads to very small uncertainties and masks any sampling issues that might have emerged as large uncertainties in P or ρ from the independent GEMC simulations. Thus, for GEMC simulations that calculate Z_i (as is our recommendation), the value of Z as a metric is mostly limited to detecting molecular weight issues and typos/scripting errors when data are prepared for publication or transferred to databases. It should also be noted here that use of Z_i requires that the total volume is sufficiently large that the vapor phase always contains at least one particle. On balance, our recommendation for GEMC simulations is still to calculate Z as a mechanical observable (because of the shortcomings in Z_e due to the correlation between P and ρ and the issues with incorrect error propagation) and to rely on an alternative metric to look for poor sampling. That being said, calculation of Z_e is still a very good metric for GCMC-HR and liquid-slab molecular dynamics simulations because Eqn. 2 for the former and the usually small vapor region (occupied frequently by no particles at low T_r) for the latter make calculation of Z_i impractical.

As a metric, Z will only be useful to the extent that deviations can be quantified. To quantify how much deviation might be allowed in high-quality VLE simulations, we use the ratio of the observed compressibility factor at a given T_r (Z_{sim}) and the expected compressibility factor based on the average behavior of a representative group of compounds (Z_{exd}). Ratios closer to unity will signify more robust simulations. Z_{exd} is first determined by fitting a logarithmic function to a reliable data set—here we use the validation data from the TraPPE database—and then using the resulting equation to calculate Z at a given T_r . For this fit, we also include Z_{crit} obtained from the ratio of the extrapolated P_{crit} and ρ_{crit} .

As shown in Figure 5, the data from the TraPPE validation set can be well described by the following empirical equation that strikes a balance between closeness of the fit and number of adjustable parameters, while also yielding physically meaningful Z values at $T_r = 1$ and low T_r :

$$Z_{\text{exd}}(T_r) = \begin{cases} Z^* + (1 - Z^*) \left[1 - ((1 + \alpha)T_r - \alpha)^\beta \right]^\gamma & \text{for } T_r \geq T_r^* \\ 1 & \text{for } T_r < \alpha/(1 + \alpha) \end{cases} \quad (7)$$

where the two powers with values of $\beta = 1.8$ and $\gamma = 0.46$ and $\alpha = 0.9$, that controls the reduced temperature ($T_r^* = \alpha/(1 + \alpha) = 0.4737$) below which Z is always set equal to unity, are three fitting parameters. The value Z^* is the average critical compressibility for the validation compounds, and its value of 0.277 is determined before the three-parameter fit. The use of a single function for Z_{exd} ignores the fact that different molecule types will have different compressibilities as the vapor density increases and the critical point is approached. For example, $Z_{\text{crit}} = 0.2295$ and 0.307 ± 0.005 for water⁴⁴ and Lennard-Jonesium,⁴⁰ respectively, spanning the range observed for the TraPPE validation models. For a more specific data set of molecules with the same functional group or molecular shape, the coefficients could be modified to yield a tighter fit. For a mixed data set such as the TraPPE database, this best-fit function (Eqn. 7), with an R^2 coefficient of 0.993, provides a robust empirical relationship between T_r and Z_{exd} for most models.

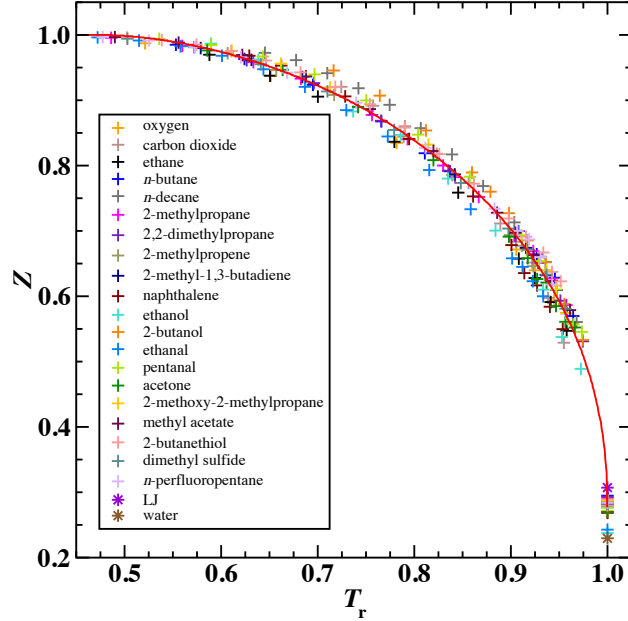


Figure 5: Compressibility factors for the TraPPE validation data. Symbols and colors for the 20 validation models are as described in the legend. Two additional critical compressibilities are shown for reference: experimental data for water⁴⁴ and the Lennard-Jones (LJ) model.⁴⁰ The red line indicates the best-fit empirical function (Eqn. 7, $R^2 = 0.993$).

Next, we calculate the ratio of Z_{sim} to Z_{exd} for all distinct Z values in the TraPPE database and plot the distribution of the natural logarithm of this ratio for the entire data set (see Figure 6). Recognizing that simulations near the critical point and those at low T_r have different challenges than those that are at intermediate T_r , the distributions of $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ are plotted for three temperature ranges: $T_r \leq 0.70$, $0.70 < T_r \leq 0.85$, and $T_r > 0.85$. Finally, a Gaussian function is fit to each distribution to provide a numerical estimate of the mean and standard deviation, from which we can quantitatively distinguish expected values from outliers. The results of these fits are shown in Table 5.

For both development and validation data, the spread of the distribution increases with increasing T_r . A wider distribution of $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ is expected near the critical point as we are using a general empirical value in place of a group-specific Z_{crit} . The range of models in the TraPPE database will therefore lead to more variation in $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ as T_r approaches unity. As Figure 6 shows, the validation data all have narrower distributions in each of the three T_r ranges when compared to the development data. For both development and

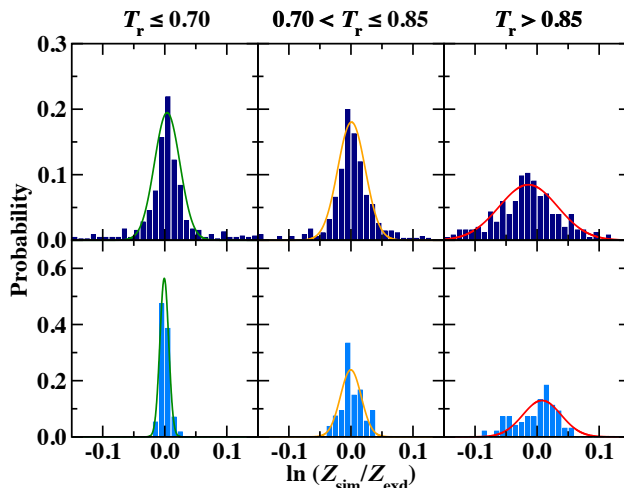


Figure 6: Gaussian fits to the distribution of $\ln(Z_{\text{sim}}/Z_{\text{extd}})$ for the TraPPE database. Development data are shown in the top row (black); validation data in the bottom row (blue).

Table 5: **Numbers of Data Points and Outliers, Mean and Standard Deviation for $\ln(Z_{\text{sim}}/Z_{\text{extd}})$ Distributions of Development^a and Validation Data**

Range	Dataset	N_{data}	N_{out}	$\bar{x}$	σ
$T_r < 0.70$	dev	466	139	0.004	0.020
	val	57	—	−0.001	0.007
$0.70 < T_r < 0.85$	dev	351	52	0.001	0.022
	val	54	—	−0.001	0.017
$T_r > 0.85$	dev	255	38	−0.014	0.047
	val	98	—	0.008	0.031

^aFor the determination of $\bar{x}$ and σ for the development set, the four molecules with known incorrect molecular weight were excluded; i.e., 14, 12, and 6 data points were removed for the low, intermediate, and high T_r ranges, respectively. However, these points are included in the count of outliers.

validation data, the mean is near zero illustrating that Eqn. 7 yields a good fit. Though the validation dataset is small, given the higher standards used for these simulations and the overarching goal of producing a metric that sets a very high bar for success, we use the narrower validation distributions to determine the acceptable range for our Z metric. For each T_r region, the acceptable range is taken as $\bar{x} \pm 3\sigma$, which should contain 99.7% of well-behaved compressibility factors. When $T_r \leq 0.70$, $\ln(Z_{\text{sim}}/Z_{\text{extd}})$ should fall between −0.03 and 0.03, for $0.70 < T_r \leq 0.85$, the acceptable range is −0.06 and 0.05, and at $T_r > 0.85$, $\ln(Z_{\text{sim}}/Z_{\text{extd}})$ should be between −0.09 and 0.11.

Table 6: **Selected** $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ **Outliers**^a

Molecule	T_r	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	NB
ethene	0.80	-0.15_4	-0.06
ethylbenzene	0.72	0.18	0.05
	0.80	0.11	0.05
<i>p</i> -xylene	0.64	-0.16	-0.03
<i>o</i> -xylene	0.63	0.14	0.03
	0.94	0.15	0.11
ethanal	0.56	-0.36	-0.03
propan-2-ol	0.60	-0.35	-0.03
octan-1-ol	0.48	0.15	0.03
nitrobenzene	0.50	-0.53	-0.03
pentanethiol	0.50	1.17	0.03
	0.53	0.62	0.03
	0.56	0.35	0.03
anthracene	0.62	-0.5_3	-0.03
	0.70	-0.4_2	-0.03
	0.78	-0.5_2	-0.06
	0.89	-0.5_3	-0.09

^a When statistical uncertainties are reported in the original publication for P_{vap} and ρ_{vap} , propagated errors for $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ have been included as subscripts. The nearest boundary (NB) that each outlier exceeds is also shown for easy reference.

Applying the metric, we find that there are many data points that emerge as Z outliers in the TraPPE database. For the 159 TraPPE models with both VLCC and P_{vap} data, 95 (60%) have at least one Z outlier. This corresponds to 229 individual data points (21% of the total data available). A complete listing of Z outliers is provided in the Supporting Information (Table S3). As illustrative examples of the types of outliers observed here, Table 6 provides numerical data for several compounds. Data for ethene, ethylbenzene, *p*-xylene, *o*-xylene, propan-2-ol, and nitrobenzene (Figure 4, Panel D: black circles, red diamonds, magenta diamonds, and orange diamonds, respectively; Panel E: magenta circle; Panel G: blue circle) are all suspected typos or data entry errors from TraPPE-4, TraPPE-5, and TraPPE-7.^{15,16,18} In each case, there are one or two Z values that are obviously outside the curve created by the other data points. Ethanal¹⁷ (Figure 4, Panel F, red diamonds) and pentanethiol¹⁹ (Figure 4, Panel J, purple circles) are both examples of data from GCMC-

HR simulations. Here, the lowest T_r is a very significant outlier for both: $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ for ethanal’s lowest T_r exceeds 50σ and pentanethiol’s low- T $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ exceeds 150σ . Pentanethiol also shows a rapid, exponential-like decrease in Z as the temperature increases, until eventually the data becomes well-behaved nearer to the critical point. As discussed above, the histogram-reweighting technique distributes errors associated with the lowest temperatures into the nearby higher temperatures creating this distinctive curvature, seen most characteristically for pentanethiol. There are, of course, also sampling problems in the development data obtained from GEMC simulations, but these are not as apparent and usually also involve large uncertainties for Z_e . One example is octan-1-ol (Figure 4, Panel E: orange circle) with $Z_e \approx 1.2$. Under-sampling of the vapor phase at lower T_r can also be seen in the deviations from linearity in the Clausius-Clapeyron plots for both GCMC-HR and many GEMC examples (see next section). The last of the selected outliers in Table 6 is anthracene (Figure 4, Panel L, red squares),²⁵ which is one of four molecules that used the wrong molecular weight to compute the VLCC during transferability tests of models that were parameterized to (usually) compounds of lower molecular weight or fewer substituent groups. Discerning the actual shape of the anthracene curve is difficult due to the large error bars, but the Z values appear to have the right curvature, just offset by a consistent factor of about 1.6 (equivalent to the molecular weight ratio of C_{28} and $\text{C}_{14}\text{H}_{10}$). This leads to $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ values for anthracene close to -0.5 at each temperature.

A Metric Based on Clausius-Clapeyron Triads

As discussed in the previous section, the Z metric (particularly, when Z_i is used) can fail to detect inadequate sampling in GEMC simulations. Here, we propose a complementary metric that evaluates the deviation of the vapor pressure from expected behavior. The Clausius-Clapeyron (CC) equation can be used to predict the vapor pressure at T_i when the

enthalpy of vaporization, ΔH_{vap} , and vapor pressure at T_j are known:

$$\ln \left(\frac{P(T_i)}{P(T_j)} \right) = -\frac{\Delta H_{\text{vap}}}{R} \left(\frac{1}{T_i} - \frac{1}{T_j} \right) \quad (8)$$

Thus, data for $\ln P$ plotted against inverse temperature should fall onto a straight line with a slope of $-\Delta H_{\text{vap}}/R$:

$$\ln P = \left(-\frac{\Delta H_{\text{vap}}}{R} \right) \frac{1}{T} + C \quad (9)$$

Such CC plots are shown in numerous TraPPE publications and other studies of VLE. However, the CC equation is derived from the Clapeyron equation which equates the local slope of the coexistence line to the entropy or enthalpy and the volume change of the transition:

$$\left(\frac{dP}{dT} \right)_{\text{VLCC}} = \frac{\Delta S_{\text{vap}}(T)}{(\bar{V}_{\text{vap}}(T) - \bar{V}_{\text{liq}}(T))} = \frac{\Delta H_{\text{vap}}(T)}{T(\bar{V}_{\text{vap}}(T) - \bar{V}_{\text{liq}}(T))} \quad (10)$$

The derivation of the CC equation is based on a set of assumptions that are not valid for the entire vapor–liquid coexistence region. Specifically, it is assumed that $\bar{V}_{\text{vap}} \gg \bar{V}_{\text{liq}}$ and $\bar{V}_{\text{vap}} = RT/P_{\text{vap}}$ (or $Z_{\text{liq}} = 0$ and $Z_{\text{vap}} = 1$, respectively) and that ΔH_{vap} is constant.

At higher T_r , ΔH_{vap} changes rapidly and approaches zero while the volume difference deviates strongly from the ideal gas law and also approaches zero.⁴⁵ The approximate linearity of the CC plot is only due to cancellation of these errors. Interestingly, the deviation from CC behavior can be larger at low T_r , despite the fact that the first two approximations hold, because ΔH_{vap} decreases with increasing temperature and there is no error cancellation, i.e., $(dP/dT)_{\text{VLCC}}$ increases with decreasing T and there is downward curvature in the CC plot.

In principle, ΔH_{vap} is available from GEMC simulations for vapor–liquid equilibria, but it is less often reported. In addition, ΔH_{vap} is difficult to evaluate from molecular dynamics simulations of an interfacial system. Thus, to ensure wider applicability, the complementary pressure metric should not rely on knowledge of ΔH_{vap} , meaning the Clapeyron equation cannot be used. Although one could use Eqn. 9 to fit data over the entire temperature range,

this would not lead to accurate values of the expected vapor pressure, P_{exd} , for most of the temperatures because of the downward curvature in the CC plot. Linearity for the entire coexistence dataset, as measured by R^2 , would be too coarse a metric. Instead, we utilize triads of neighboring temperatures for our P -based metric, where two of these temperatures are used to determine the local slope, $-\Delta H_{\text{vap}}/R$, via Eqn. 8, and P_{CC} for the third point is determined by also applying the CC equation.

The P -based metric is then developed in a manner similar to the Z -based metric, but with the added understanding that the expected P_{exd} can deviate somewhat from P_{CC} . That is, we create distributions of $\ln(P_{\text{sim}}/P_{\text{CC}})$, then use the mean to determine expected values, $\ln(P_{\text{CC}}/P_{\text{exd}})$, and let the standard deviation establish numeric boundaries, all based on high-quality simulations (i.e., the validation data). Since $\ln(P_{\text{CC}}/P_{\text{exd}})$ can deviate from zero, this approach includes an expectation of curvature in the CC plot, i.e., $\ln(P_{\text{sim}}/P_{\text{exd}}) = \ln(P_{\text{sim}}/P_{\text{CC}}) + \ln(P_{\text{CC}}/P_{\text{exd}})$. To find P_{CC} , we use two of the points in each triad to predict the P of the third. Predictions are done in three ways: CC_{LOW} , where the two higher T_{r} points of the triad are used to predict P_{CC} at the lowest T_{r} of the triad; CC_{HIGH} , where the two lower T_{r} points are used to predict P_{CC} at the highest T_{r} ; and CC_{MID} , where the highest and lowest T_{r} points are used to predict P_{CC} at the middle T_{r} . If data at N_T temperatures are available, then there are $N_T - 2$ CC_{LOW} , $N_T - 2$ CC_{MID} , and $N_T - 2$ CC_{HIGH} triads. As an example, the CC plot for propan-1-ol¹⁶ is shown in Figure 7. For CC_{LOW} and CC_{HIGH} , the predicted point is extrapolated outside of the two points used in the linear fit. The range of this extrapolation can be determined by taking the ratio of two differences, the difference between the inverse temperatures of the predicted point and its nearest neighbor divided by the difference between the two inverse temperatures used to make the prediction. Ideally, this ratio should be close to 1.0 (i.e., equally spaced inverse temperatures). Larger extrapolation ranges are more likely to lead to inaccurate values of P_{exd} . As an aside, when the Gibbs–Duhem integration approach⁴⁶ is used to trace the coexistence line, then analysis of ΔH_{vap} triads would likely be more useful than of the CC triads.

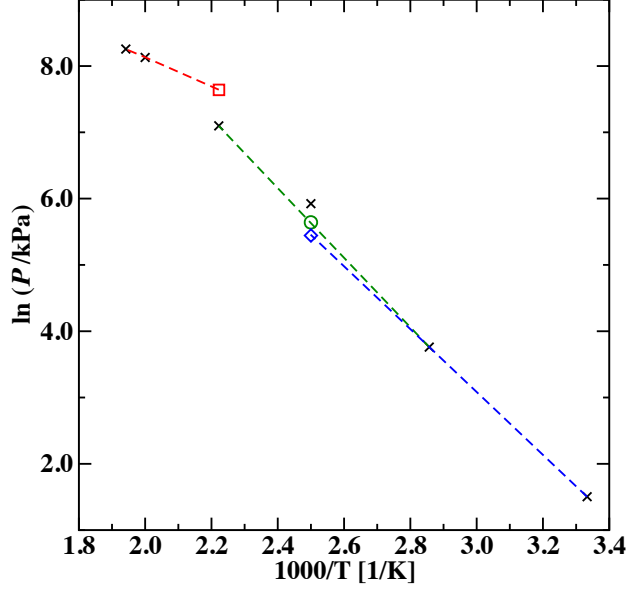


Figure 7: Clausius-Clapeyron triads obtained from simulations of the TraPPE-UA model for propan-1-ol.¹⁶ Original simulation data are shown as black crosses. Linear CC fits for CC_{LOW} , CC_{HIGH} , and CC_{MID} are shown as red, blue, and green dashed lines, and the corresponding P_{ext} values are depicted as red square, blue diamond, and green circle, respectively.

Analysis of CC Triads from the TraPPE Database

The use of triads reduces the size of our overall TraPPE dataset compared to the Z metric, from 1072 total points to 754 (see Table 2), and leads to three distinct sets of P_{ext} values, where a P_{sim} data point at a specific T_r is compared to a P_{ext} value one to three times. The P_{sim} at the highest T_{sim} will only be utilized once as the highest T in a triad (similarly just once for the lowest T_{sim} as the lowest T in a triad), the next points inward can be predicted twice (CC_{HIGH} or CC_{LOW} and CC_{MID}), while interior points will be involved in three triads. Of course, an erroneous P_{sim} at a single T_r can lead to outliers for all three triads involving this single T_r ; i.e., the total number of $\ln(P_{\text{sim}}/P_{\text{ext}})$ outliers is larger than the number of erroneous pressure values.

As with the Z -based metric, provided with a reliable value for P_{ext} , the amount of tolerable deviation in P_{sim} can be determined from our validation data. We again consider three T_r ranges, differing in their relative nearness to the critical point, and fit Gaussian

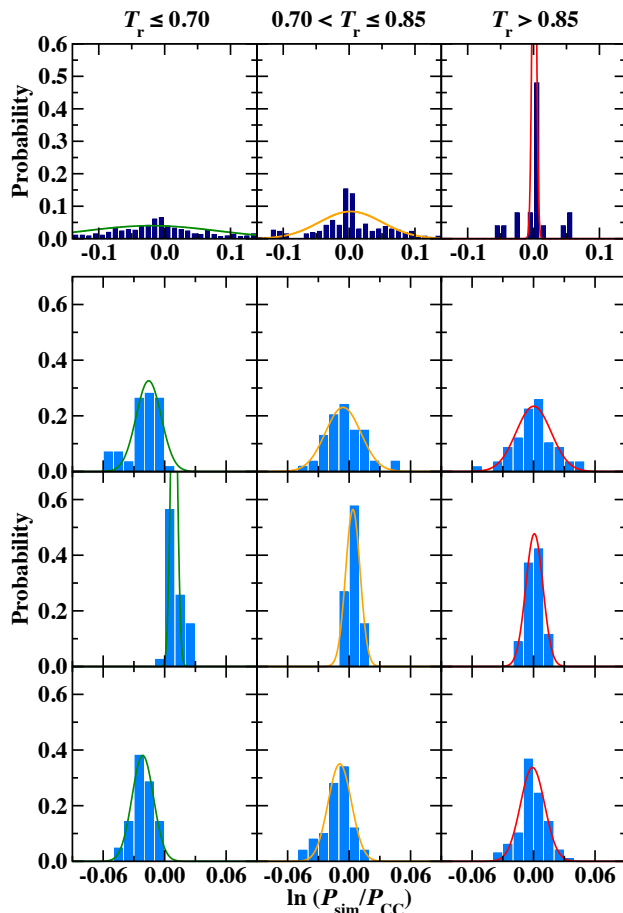


Figure 8: Gaussian fits to the distributions of $\ln(P_{\text{sim}}/P_{\text{CC}})$ for the TraPPE database. Development data are shown in the top row (dark blue) for the CC_{LOW} predictions. The following three rows show validation data (light blue) for the CC_{LOW} , CC_{MID} , and CC_{HIGH} predictions, respectively.

functions separately for each. The results of these fits are shown in Figure 8 and Table 7. For the development data, the standard deviation is on average ten times larger in magnitude than the value of $\bar{x} = \ln(P_{\text{CC}}/P_{\text{exd}})$, suggesting there is no statistically significant deviation from linearity in the CC plot. The distributions for the validation set are much narrower than those for the development set (by factors of 11, 4, and 2 for $T_r \leq 0.70$, $0.70 < T_r \leq 0.85$, and $T_r > 0.85$, respectively), with the exception of the high- T_r range for CC_{LOW} . We attribute this to the validation effort emphasizing precise and accurate estimation of the critical point and including a much larger number of data points with $T_r > 0.9$. The average value of the third highest T_r (highest T_r considered for CC_{LOW}) for the 159 development compounds is

Table 7: **Numbers of Data Points and Outliers, Mean ($\bar{x} = \ln(P_{CC}/P_{\text{exd}})$), and Standard Deviation for the $\ln(P_{\text{sim}}/P_{CC})$ Distributions of the Development and Validation Data^a**

Triad	Dataset	$T_r \leq 0.70$				$0.70 < T_r \leq 0.85$				$T_r > 0.85$			
		N_{data}	N_{out}	$\bar{x}$	σ	N_{data}	N_{out}	$\bar{x}$	σ	N_{data}	N_{out}	$\bar{x}$	σ
CC _{LOW}	dev	462	236	-0.020	0.100	267	107	0.002	0.048	25	3	0.001	0.003
	val	57	–	-0.016	0.012	54	–	-0.006	0.017	58	–	0.001	0.017
CC _{MID}	dev	309	206	0.006	0.050	341	165	0.003	0.032	104	29	-0.003	0.018
	val	39	–	0.009	0.003	52	–	0.004	0.007	78	–	-0.001	0.008
CC _{HIGH}	dev	175	85	-0.008	0.086	324	170	-0.013	0.063	255	124	0.004	0.053
	val	21	–	-0.021	0.010	50	–	-0.009	0.011	98	–	-0.001	0.012

^a The T_r of the predicted point determines the assignment to a specific T_r range, i.e., the two other points in the triad may not fall into this range.

0.88, whereas it is 0.93 for the validation data. Furthermore, the shape of the $\ln(P_{\text{sim}}/P_{CC})$ distribution for the development data at $T_r > 0.85$ is also not well described by the Gaussian curve, and the width may be somewhat underestimated. Considering the three CC triads, the number of data points used in the $\ln(P_{\text{sim}}/P_{CC})$ fitting is largest near the critical point for the validation data, but the opposite holds for the development set.

With the smaller widths of the distributions, the validation data show distinct deviations from linearity in the CC plot as indicated by $\ln(P_{CC}/P_{\text{exd}})$ that differ significantly from zero. As expected from the increase in $(dP/dT)_{\text{VLCC}}$ with decreasing T (i.e., downward curvature in the CC plot), $\ln(P_{CC}/P_{\text{exd}})$ values are negative, positive, and negative for CC_{HIGH}, CC_{MID}, and CC_{LOW}, respectively. The average deviation from linearity is largest for $T_r \leq 0.70$, where it is on average 2.1 times larger than the standard deviation. However, for the validation data, the distributions for all three CC metrics at $T_r \leq 0.70$ are not well described by the Gaussian fits with clear tails on the side further away from zero. There are two potential reasons for the skew: (i) some of the data fall below $T_r \leq 0.55$, and it could be that data at these lower T_r require a larger magnitude of $\bar{x}$, but we do not have enough data to establish separate distributions for $T_r \leq 0.55$; (ii) the simulations for most of the validation compounds use constant temperature steps that lead to increasing step sizes in the inverse temperature used for the CC metric and, as a result, the extrapolation range increases with decreasing

temperature.

Shifts in the distribution are still visible for $0.70 < T_r \leq 0.85$, but here the standard deviation is on average 1.9 times larger than the shift. For $T_r > 0.85$, we do not find any significant shift presumably due to the cancellation of errors resulting from the assumptions made in the derivation of the CC equation. Whereas different values of Z_{crit} (due to hydrogen bonding and molecular shape) lead to a significant increase (by a factor of 4) of the width of the $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ distribution for $T_r > 0.85$, the increase in the width of the $\ln(P_{\text{sim}}/P_{\text{exd}})$ distributions is only a factor of 1.7; that is, the P metric holds well for different molecule types. Finally, the $\ln(P_{\text{sim}}/P_{\text{CC}})$ distribution (see Figure 7) is very broad for the development data at $T_r \leq 0.70$, this indicates that there are indeed sampling challenges that lead to problems for the low- T_r simulations.

Identification of $\ln(P_{\text{sim}}/P_{\text{exd}})$ outliers for the CC_{LOW} , CC_{MID} , and CC_{HIGH} tests results in high redundancy though there are differences in the specific points designated as outliers. However, problematic simulations will produce outliers for each test. For simplicity and because we expect the lower T_r to be most informative for our metric (i.e., they are most likely to suffer from sampling problems due to low acceptance rates for molecule transfers and low numbers of molecules in the vapor phase), we focus primarily on the CC_{LOW} test in further analysis here, but all of the data for each CC_{triad} test are available in the Supporting Information (see Table S4).

As was done for the Z metric, the acceptable range for $\ln(P_{\text{sim}}/P_{\text{CC}})$ is given as $\bar{x} \pm 3\sigma$ (i.e., equivalent to the acceptable range for $\ln(P_{\text{sim}}/P_{\text{exd}})$ being $0 \pm 3\sigma$) for the two higher T_r regions, but $\bar{x} \pm 6\sigma$ is used for $T_r \leq 0.7$ to account for the skew in the data. Specifically for the CC_{LOW} test, when $T_r \leq 0.70$, $\ln(P_{\text{sim}}/P_{\text{CC}})$ should be between -0.09 and 0.06 , for $0.70 < T_r \leq 0.85$, the acceptable range is -0.06 and 0.05 , and at $T_r > 0.85$, $\ln(P_{\text{sim}}/P_{\text{CC}})$ should be between -0.06 and 0.06 . The boundaries of the acceptable ranges for CC_{MID} and CC_{HIGH} are provided in the caption of Table S5.

Applying the CC_{LOW} test to the TraPPE database for the development molecules results

in 140 (88%) of the molecules having at least one P_{sim} that falls outside of the acceptable range. The outliers comprise 346 individual data points, or 46% of the CC_{LOW} triads. For comparison, the CC_{MID} and CC_{HIGH} tests lead to 400 and 379 outliers. Thus, extrapolation to low and high temperatures and interpolation to an intermediate temperature yield rather similar number of outliers but, again, the validation data lead to a narrower distribution for interpolation (CC_{MID}) and, in turn, narrow bounds.

Most of the CC_{LOW} outliers come from the lowest three temperatures simulated for a given model. The numbers of CC_{triad} outliers are about a factor of 1.5 to 1.7 higher than the 229 outliers found with the Z metric. At first glance, the boundaries used for the CC_{triad} tests appear to be more restrictive than those for the Z metric. However, as discussed previously, calculation of P using the virial equation (see Eqn. 1) in GEMC simulations leads to strong correlation between P and ρ_{vap} at low T_{r} and, as a result, masks problematic simulations in the Z metric. A second reason is that more temperatures are used for the validation set (slightly above 10 on average per compound versus less than 7 for the development set), such that the T_{r} intervals are on average larger for the development set. This increases the extrapolation range and the expected deviation from CC behavior. Thus, the larger inverse temperature intervals or sometimes also uneven intervals in absolute temperature may cause some false positives in the development data. In principle, it would be advantageous to make the boundaries of the CC_{triad} tests also a function of the extrapolation range, but we do not have sufficient validation data to do this in a statistically meaningful manner, and it would introduce further complexity. Thirdly, differences in hydrogen bonding ability and molecular shape result in a larger spread for the Z metric (leading to a more permissive Z metric), whereas the curvature of the CC plots may be more universal.⁴¹

A set of selected CC_{LOW} outliers is depicted as CC plots in Figure 9 and some numerical values are provided in Table 8. The lowest- T_{r} vapor pressures of quinoline and pyridazine are the most extreme outliers in the CC_{LOW} test, and the “hockey-stick” shape of their CC curves leads to deviations of about 250 and 200 σ away from $\bar{x}$. These data correspond to $T_{\text{r}} = 0.37$

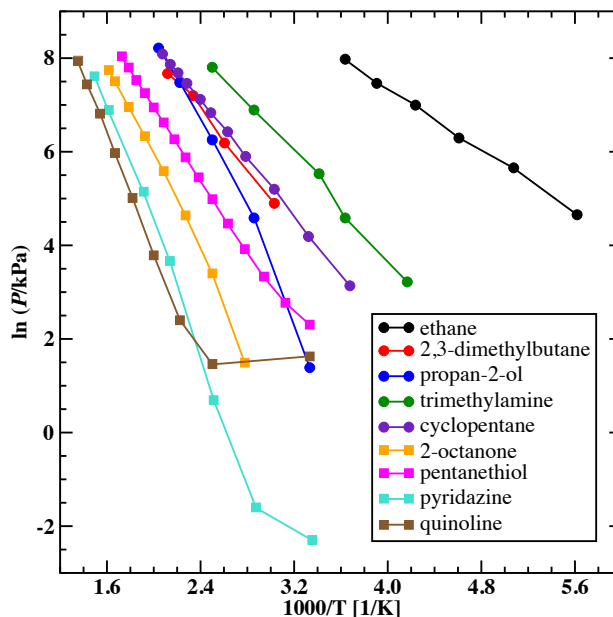


Figure 9: Clausius-Clapeyron plots for selected TraPPE models. Colors and symbols are as described in the legends. Connecting lines are provided only as an aid to visualization.

and 0.39, that are some of the lowest T_r attempted in the TraPPE database. These were used to compare to experimental liquid densities at near-ambient conditions, and the P data at these temperatures are not shown in the CC plots included in the respective publications but only appear in the respective Supporting Information.^{24,25} Given the significant deviation from the remainders of the CC curves, the GEMC simulations at these temperature were likely not successful, and the data made it into the supporting information only due to scripting errors.

Ethane,¹¹ 2,3-dimethylbutane,¹² and propan-1-ol,¹⁶ (see Figure 7) are examples of early TraPPE development efforts when properties were calculated from block averages in a single simulation trajectory rather than truly independent simulations. At the time, priority was given to fitting ρ_{liq} and T_{crit} with high accuracy (less than 1% error), while lower targets were accepted for the vapor phase. The CC plots show small but clear deviations from linearity for these models, often over the entire T_r range, and the sensitive CC_{triad} metric flags them as outliers. Trimethylamine¹⁸ and cyclopentane²² are more recent examples, but involve CBMC sampling challenges as a highly branched EH model or as a ring polymer, respectively. Of

Table 8: **Selected CC_{LOW} Outliers with Their Nearest Boundary (NB_{CC}), Extrapolation range (ER), and corresponding $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ and NB_Z Values**

Molecule	T_r	$\ln(P_{\text{sim}}/P_{\text{CC}})$	NB _{CC}	ER	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	NB _Z
ethane	0.59	-0.26	-0.09	1.16	0.01 ₅	0.03
2,3-dimethylbutane	0.76	-0.43	-0.06	1.23	0.00 ₉	0.05
propan-1-ol	0.84	-0.55	-0.06	3.82	-0.03 ₈	-0.06
propan-2-ol	0.60	-0.97	-0.09	1.33	-0.35 ₁₄	-0.03
cyclopentane	0.64	0.136	0.06	1.59	0.00 ₃	0.03
trimethylamine	0.55	0.88	0.06	2.37	0.01 ₆	0.03
2-octanone	0.56	-0.38	-0.09	1.22	-0.54	-0.03
pentanethiol	0.50	0.16	0.06	1.13	1.17	0.03
pyridazine	0.39	2.38	0.06	1.34	0.5 ₆	0.03
quinoline	0.37	2.99	0.06	3.00	0.04 ₂	-0.03

particular interest in the pursuit of metrics for simulation quality is the fact that none of these five compounds shows an outlier in the Z metric (see Table 8). Furthermore, quinoline at $T_r = 0.37$ yields $Z = 0.964 \pm 0.022$, a much less alarming value than the huge deviation found by the CC_{triad} metric. Conversely, an error with assigning molecular weights would be entirely invisible for the CC_{triad} test because the pressure is calculated from the forces and is molecular-weight independent. This clearly shows that the two metrics are complementary, and together give a fuller picture of simulation quality.

A smaller number of TraPPE models did result in both Z and CC_{triad} outliers and, generally, the failures are large in these cases for both metrics. Propan-2-ol¹⁶ suffers from sampling challenges due to hydrogen-bonding at the lowest T_r (and also a large temperature step down to this T_r), but may also have a data entry mistake. This data point for propan-2-ol is the outlying magenta circle in Panel E of Figure 4. Pentanethiol¹⁹ and 2-octanone¹⁷ are among the examples from GCMC-HR simulations that are insufficiently sampled at low T_r . This is reflected as large upward or downward curvature (instead of scatter) in the CC plots (see Figure 9), as well as the corresponding curvatures in the Z versus T_r plots (purple circles in Panel J for pentanethiol and green circles in Panel F for 2-octanone in Figure 4). Many of the EH models for aromatics^{24,25} also include significant outliers in the CC_{triad} metric, but not as severe as quinoline and pyridazine; transfer moves for these large, rigid molecules are

particularly challenging at low T_r .

Conclusions

To summarize, the analysis of the TraPPE database with the highlighted examples provided here helps to demonstrate the importance and power of the Z and CC_{triad} based metrics for simulation quality. While Z is a good metric for detecting many types of simulation errors (insufficient sampling in MD with interfacial boxes and GCMC-HR simulations, molecular weight discrepancies, and typos), it routinely fails to account for insufficient sampling in GEMC simulations. This is simply a consequence of the way that P is calculated in GEMC and signifies the need for a complementary metric, such as the CC_{triad} metric, for these types of simulations. A more powerful metric emerges when both Z and CC_{triad} can be evaluated together as part of the same set of simulations. Most of the pronounced errors (with deviations larger than 10σ) found in previous simulations of TraPPE models, if not all, could have been corrected with more careful attention to the simulation data; attention that in the future can be helpfully directed by the Z and CC_{triad} metrics. To this extent, a Python script for performing the Z and CC_{triad} tests is provided as Supporting Information.

For GEMC, Z should not be calculated using separate ensemble averages of P and ρ_{vap} (i.e., $Z_e = (\langle P \rangle / \langle \rho_{\text{vap}} \rangle)(W/RT)$) together with error propagation at the end of the simulation because, at low T_r , P is highly correlated with ρ_{vap} and the statistical error can be overestimated by as much as two orders of magnitude. At high T_r , Z_e underestimates the true $\langle Z \rangle$ value as instantaneous values of z and ρ_{vap} become more correlated; a histogram analysis of the instantaneous z value distributions leads to more accurate $Z_i = \langle P / \rho_{\text{vap}} \rangle (W/RT)$ (and also ρ_{liq} , ρ_{vap} , and P) ensemble averages. In other words, Z should be calculated as Z_i using histogram analysis. A Python script for carrying out the histogram analysis is provided as Supporting Information.

More persistent simulation challenges, particularly for GEMC, where Z should always

be well-behaved for properly executed simulations, require a second, complementary metric. The Clausius-Clapeyron curve (or, when available, directly calculated values of the ΔH_{vap} for sufficiently small temperature steps) can be used to detect sampling problems in GEMC simulations. Due to the inverse-temperature relation of the Clausius-Clapeyron curve, the CC_{triad} metric will be most informative for simulation quality when simulation data are available at equally spaced T^{-1} values. Satisfying both the Z and CC_{triad} metrics will require longer trajectories, larger system sizes, and/or better simulation protocols (and, certainly, more computer time). Most of the less significant outliers (with deviations falling into the range from 4 to 9σ) found for the TraPPE development data are a consequence of the more limited computer resources available at the time when the development was done, i.e., data from these earlier simulations have larger statistical uncertainties.

Taken as a whole, the analysis performed here indicates that high-quality simulation data, while certainly achievable, are by no means easy to achieve. The fact that MMS methods are increasingly accessible (and cheap) must be balanced by the much harder task of properly applying these methods to generate robust and reliable data. Still, when simulators pay careful attention to the details of the method and resist the temptation to underestimate the difficulty of performing a simulation correctly, high-quality data can be produced. The Z and CC_{triad} metrics developed here can help guard against that temptation and set the standard for high-quality VLE data for moderately complex organic molecules described by molecular mechanics force fields. When the complexity of the compounds (e.g., number of atoms, degree of branching, or hydrogen-bonding capability) increases or when more expensive descriptions of the interactions (e.g., Kohn-Sham density functional theory) are used, then lower precision and more scatter of the VLE data is to be expected.

Acknowledgement

This research is primarily supported through the National Science Foundation (CBET-1835067) and Wheaton College. The work on correlations between properties and histogram analysis is supported by the Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under Award DE-FG02-17ER16362. The development of the TraPPE database and website were supported by the National Science Foundation (RUI-1159731). Computational resources from the Minnesota Supercomputing Institute are also gratefully acknowledged. Several validation simulations were performed by Siepmann group members not included as authors (and not previously published): C. Bunner (ethanal, dimethyl sulfide, and *n*-perfluoropentane), J. Chen (2-butanethiol), E. Fetisov (2-methylpropene), D. Harwood (oxygen and 2,2-dimethylpropane), B. Liang (2-butanol), M. Shah (carbon dioxide), K. Struksheats (acetone, 2-methoxy-2-methylpropane, naphthalene, and 2-methyl-1,3-butadiene), and B. Xue (data for Table 4).

Supporting Information Available

Additional simulation details and numerical VLCC data for the validation models. Description of GitHub repositories containing the version of the MCCC-S-MN software together with input and output files used for validation simulations of pentanal, the Python script for the Z and CC_{triad} tests, and the Python script for the histogram analysis of coexistence properties from GEMC simulations. Tables listing all the $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ and $\ln(P_{\text{sim}}/P_{\text{CC}})$ outliers for the TraPPE database.

References

- (1) Kofke, D. A.; Siepmann, J. I.; Chirico, R. D.; Kroon, M.; Mathias, P. M.; Sadowski, G.; Wu, J.; Brennecke, J. F. Editorial: Molecular Modeling and Simulation in JCED. *J. Chem. Eng. Data* **2016**, *61*, 1–2.
- (2) Frenkel, D. Simulations: The Dark Side. *Eur. Phys. J. Plus* **2013**, *128*, 1–21.
- (3) van Gunsteren, W. F.; Daura, X.; Hansen, N.; Mark, A. E.; Oostenbrink, C.; Riniker, S.; Smith, L. J. Validation of Molecular Simulation: An Overview of Issues. *Angew. Chem. Int. Ed.* **2018**, *57*, 884–902.
- (4) Schappals, M.; Mecklenfeld, A.; Kröger, L.; Botan, V.; Köster, A.; Stephan, S.; Garcia, E.; Rutkai, G.; Raabe, G.; Klein, P.; Leonhard, K.; Glass, C. W.; Lenhard, J.; Vrabec, J.; Hasse, H. Round Robin Study: Molecular Simulation of Thermodynamic Properties from Models with Internal Degrees of Freedom. *J. Chem. Theo. Comput.* **2017**, *13*, 4270–4280.
- (5) Nezbeda, I.; Kolafa, J. The Use of Control Quantities in Computer Simulation Experiments: Application to the Exp-6 Potential Fluid. *Mol. Simul.* **1995**, *14*, 153–163.
- (6) Schultz, A. J.; Kofke, D. A. Quantifying Computational Effort Required for Stochastic Averages. *J. Chem. Theor. Comput.* **2014**, *10*, 5229–5234.
- (7) Nezbeda, I. Simulations of Vapor–Liquid Equilibria: Routine versus Thoroughness. *J. Chem. Eng. Data* **2016**, *61*, 3964–3969.
- (8) Panagiotopoulos, A. Z. Direct Determination of Phase Coexistence Properties of Fluids by Monte Carlo Simulation in a New Ensemble. *Mol. Phys.* **1987**, *61*, 813–826.
- (9) Panagiotopoulos, A. Z.; Quirke, N.; Stapleton, M.; Tildesley, D. J. Phase Equilibria by Simulation in the Gibbs Ensemble: Alternative Derivation, Generalization and Application to Mixture and Membrane Equilibria. *Mol. Phys.* **1988**, *63*, 527–545.

- (10) Eggimann, B. L.; Sunnarborg, A. J.; Stern, H. D.; Bliss, A. P.; Siepmann, J. I. An Online Parameter and Property Database for the TraPPE Force Field. *Mol. Simul.* **2014**, *40*, 101–105.
- (11) Martin, M. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-atom Description of *n*-Alkanes. *J. Phys. Chem. B* **1998**, *102*, 2569–2577.
- (12) Martin, M. G.; Siepmann, J. I. Novel Configurational-bias Monte Carlo Method for Branched Molecules. Transferable Potentials for Phase Equilibria. 2. United-atom Description of Branched Alkanes. *J. Phys. Chem. B* **1999**, *103*, 4508–4517.
- (13) TraPPE Website and Database Landing Page.
<http://chem-siepmann.oit.umn.edu/siepmann/trappe> (accessed September 9, 2018).
- (14) TraPPE Validation Parameters and Properties.
<http://chem-siepmann.oit.umn.edu/siepmann/trappe/validation.html> (accessed June 28, 2019).
- (15) Wick, C. D.; Martin, M. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 4. United-atom Description of Linear and Branched Alkenes and Alkylbenzenes. *J. Phys. Chem. B* **2000**, *104*, 8008–8016.
- (16) Chen, B.; Potoff, J. J.; Siepmann, J. I. 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.
- (17) 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.

- (18) Wick, C. D.; Stubbs, J. M.; Rai, N.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 7. Primary, Secondary, and Tertiary Amines, Nitroalkanes and Nitrobenzene, Nitriles, Amides, Pyridine, and Pyrimidine. *J. Phys. Chem. B* **2005**, *109*, 18794–18982.
- (19) Lubna, N.; Kamath, G.; Potoff, J. J.; Rai, N.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 8. United-atom Description for Thiols, Sulfides, Disulfides, and Thiophene. *J. Phys. Chem. B* **2005**, *109*, 24100–24107.
- (20) Lee, J.-S.; Wick, C.D.; Stubbs, J.M.; Siepmann, J.I. Simulating the Vapour–Liquid Equilibria of Large Cyclic Alkanes. *Mol. Phys.* **2005**, *103*, 99–104.
- (21) Maerzke, K.A.; Schultz, N.E.; Ross, R.B.; Siepmann, J.I. TraPPE–UA Force Field for Acrylates and Monte Carlo Simulations for Their Mixtures with Alkanes and Alcohols. *J. Phys. Chem. B* **2009**, *113*, 6415–6425.
- (22) Keasler, S. J.; Charan, S. M.; Wick, C. D.; Economou, I. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. United-atom Description of Five- and Six-membered Cyclic Alkanes and Ethers. *J. Phys. Chem. B* **2012**, *116*, 11234–11246.
- (23) Chen, B.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 3. Explicit-hydrogen Description of Normal Alkanes. *J. Phys. Chem. B* **1999**, *103*, 5370–5379.
- (24) Rai, N.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 9. Explicit-hydrogen Description of Benzene and Five-membered and Six-membered Heterocyclic Aromatic Compounds. *J. Phys. Chem. B* **2007**, *111*, 10790–10799.
- (25) Rai, N.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 10. Explicit-hydrogen Description of Substituted Benzenes and Polycyclic Aromatic Compounds. *J. Phys. Chem. B* **2013**, *117*, 273–288.

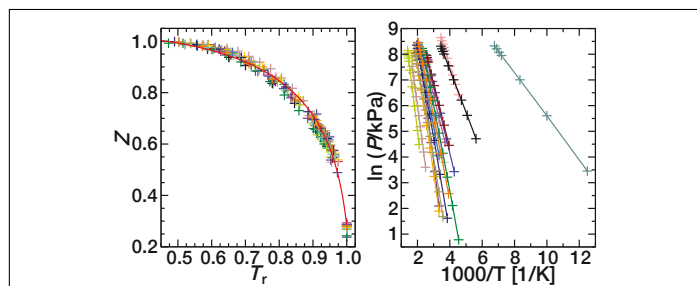
- (26) Potoff, J. J.; Siepmann, J. I. Vapor–Liquid Equilibria of Mixtures Containing Alkanes, Carbon Dioxide, and Nitrogen. *AIChE J.* **2001**, *47*, 1676–1682.
- (27) Zhang, L.; Siepmann, J. I. Direct Calculation of Henry’s Law Constants from Gibbs Ensemble Monte Carlo Simulations: Nitrogen, Oxygen, Carbon Dioxide, and Methane in Ethanol. *Theor. Chem. Acc.* **2006**, *115*, 391–397.
- (28) Ketko, M.-B. H.; Rafferty, J. L.; Siepmann, J. I.; Potoff, J. J. Development of the TraPPE–UA force Field for Ethylene Oxide. *Fluid Phase Equil.* **2008**, *274*, 44–49.
- (29) Zhang, L.; Siepmann, J. I. Development of the TraPPE Force Field for Ammonia. *Collect. Czech. Chem. Commun.* **2010**, *75*, 577–591.
- (30) Shah, M. S.; Tsapatsis, M.; Siepmann, J. I. Development of the Transferable Potentials for Phase Equilibria Model for Hydrogen Sulfide. *J. Phys. Chem. B* **2015**, *119*, 7041–7052.
- (31) Chen, B.; Xing, J.; Siepmann, J. I. Development of Polarizable Water Force Fields for Phase Equilibrium Calculations. *J. Phys. Chem. B* **2000**, *104*, 2391–2401.
- (32) Chen, B. *Phase Equilibria of Hydrocarbons, Alcohols, Water, and Their Mixtures*. Ph.D. Thesis (University of Minnesota, Minneapolis, 2001).
- (33) Maerzke, K. A.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. Coarse-grain Description for Linear Alkanes. *J. Phys. Chem. B* **2011**, *115*, 3452–3465.
- (34) Shah, M. S.; Tsapatsis, M.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. Improved United-atom Description of Ethane and Ethylene. *AIChE J.* **2017**, *63*, 5098–5110.
- (35) Bai, P.; Tsapatsis, M.; Siepmann, J. I. TraPPE–zeo: Transferable Potentials for Phase Equilibria Force Field for All-silica Zeolites. *J. Phys. Chem.* **2013**, *117*, 24375–24387.

- (36) Potoff, J. J.; Panagiotopoulos, A. Z. Critical Point and Phase Behavior of the Pure Fluid and a Lennard-Jones Mixture. *J. Chem. Phys.* **1998**, *109*, 10914–10920.
- (37) Potoff, J. J.; Errington, J. R.; Panagiotopoulos, A. Z. Molecular Simulation of Phase Equilibria for Mixtures of Polar and Non-polar Components. *Mol. Phys.* **1999**, *97*, 1073–1083.
- (38) Siepmann, J. I.; Martin, M. G.; Chen, B.; Wick, C. D.; Stubbs, J. M.; Potoff, J. J.; Eggimann, B. L.; McGrath, M. J.; Zhao, X. S.; Anderson, K. E.; Rafferty, J. L.; Rai, N.; Maerzke, K. A.; Keasler, S. J.; Bai, P.; Fetisov, E. O.; Shah, M. S.; Chen, Q. P.; DeJaco, R. F.; Chen, J. L.; Bai, X. *Monte Carlo for Complex Chemical Systems—Minnesota*, version 17.1; University of Minnesota: Minneapolis, MN, 2017.
- (39) Cortes-Morales, A. D.; Economou, I. G.; Peters, C. J.; Siepmann, J. I. Influence of Simulation Protocols on the Efficiency of Gibbs Ensemble Monte Carlo Simulations. *Mol. Simul.* **2013**, *39*, 1135–1142.
- (40) Dinpajoo, M.; Bai, P.; Allan, D. A.; Siepmann, J. I. Accurate and Precise Determination of Critical Properties from Gibbs Ensemble Monte Carlo Simulations. *J. Chem. Phys.* **2015**, *143*, 114113.
- (41) O’Connell, J. P. Vapor-phase Nonideality: When to Worry. *Chem. Eng. Prog.* **2019**, *115*(5), 63–71.
- (42) Hinkley, D. V.; On the Ratio of Two Correlated Normal Random Variables. *Biometrika* **1969**, *56*, 635–639.
- (43) Smit, B.; De Smedt, Ph.; Frenkel, D. Computer Simulation in Gibbs Ensemble. *Mol. Phys.* **1989**, *68*, 931–950.
- (44) NIST Webbook Entry for Water.

<https://webbook.nist.gov/cgi/inchi/InChI%3D1S/H2O/h1H2> (accessed June 28, 2019).

- (45) Panagiotopoulos, A. Z. *Essential Thermodynamics*; Drios Press: Princeton, 2011.
- (46) Kofke, D. A. Gibbs–Duhem Integration: A New Method for Direct Evaluation of Phase Coexistence by Molecular Simulation. *Mol. Phys.* **1993**, 78, 1331–1336.

Graphical TOC Entry



Supporting Information:

**Assessing the quality of simulations for
vapor–liquid equilibria: An analysis of the
TraPPE database**

Becky L. Eggimann,^{*,†} Yangzesheng Sun,[‡] Robert F. DeJaco,^{‡,¶} Ramanish
Singh,^{‡,¶} Muhammad Ahsan,[‡] Tyler R. Josephson,[‡] and J. Ilja Siepmann^{*,‡,¶}

[†]*Department of Chemistry, Wheaton College, Wheaton, IL, 60187-5501, United States*

[‡]*Department of Chemistry and Chemical Theory Center,
University of Minnesota, Minneapolis, MN, 55455-0431, United States*

[¶]*Department of Chemical Engineering and Materials Science,
University of Minnesota, Minneapolis, MN, 55455-0431, United States*

E-mail: becky.eggimann@wheaton.edu; siepmann@umn.edu

Contents

Glossary of compressibility and pressure variables [page 3]

Additional simulation details and numerical VLCC data for the validation models
[pages 4–13]

Description of GitHub repositories containing the version of the MCCCS–MN software together with input and output files used for validation simulations of pentanal, the Python script for the Z and CC_{triad} tests, and the Python script for the histogram analysis of coexistence properties from GEMC simulations near T_{crit} [page 14]

Tables listing all of the $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ and $\ln(P_{\text{sim}}/P_{\text{CC}})$ outliers for the TraPPE database
[pages 15–65]

Glossary of Compressibility and Pressure Variables

Z	generic compressibility factor measured or calculated through various means
Z_{crit}	compressibility factor at the critical point
Z^*	average critical compressibility factor of 20 validation compounds
Z_e	compressibility factor calculated from ensemble averages of vapor pressure and vapor density (see eq. 4)
Z_i	compressibility factor calculated from ensemble average of instantaneous vapor pressure over vapor density ratios (see eq. 5)
z	instantaneous compressibility factor for a given configuration
Z_{sim}	average compressibility factor (either Z_e or Z_i) obtained from simulation
Z_{exd}	expected compressibility factor determined from eq. 7
P	generic vapor pressure measured or calculated through various means
P_{crit}	pressure at the critical point
P_{ideal}	ideal gas pressure
P_{GEMC}	average vapor pressure calculated from the virial equation for the vapor box in a GEMC simulation (see eq. 1)
P_{GCMC}	vapor pressure calculated from the partition function in GCMC-HW simulations
P_{sim}	average vapor pressure obtained from a simulation
P_{CC}	vapor pressure of the third point in a triad determined by applying the Clausius–Clapeyron equation (eq. 8) using the enthalpy of vaporization and vapor pressure obtained from the other two points
P_{exd}	expected vapor pressure determined from $\bar{x} = \ln(P_{\text{CC}}/P_{\text{exd}})$

Validation Simulations

Additional Simulation Details

As described in the main text, all validation simulations are done in the NVT -Gibbs ensemble with coupled-decoupled configurational-bias Monte Carlo. Lennard-Jones interactions are treated with a spherical potential cutoff (r_{cut}) and analytic tail corrections. When charges are present, Coulomb interactions are calculated using the Ewald summation technique with the same r_{cut} values for the direct space part and $\kappa = 2\pi/\text{boxlength}$. For the liquid phase, the standard TraPPE value $r_{\text{cut}} = 14 \text{ \AA}$ is used for all simulations with linear dimensions typically falling into the 30 to 50 $\text{\AA}$ range. To yield predicted saturated vapor pressure and density with the desired precision requires vapor phases containing on average 10-20% of the molecules and, due to the exponential dependence of vapor pressure on temperature, a much larger spread of volumes. In particular at low reduced temperatures with very large vapor volumes, the r_{cut} value is adjusted to be approximately 40% of the boxlength to yield a good balance for the computational cost of the direct and reciprocal space parts of the Ewald summation.

At high reduced temperatures near the critical point, Gaussian fits to property histograms are used to determine ρ_{liq} , ρ_{vap} , P_{vap} , and Z_{vap} . Critical properties are obtained using VLCC data from temperatures that exceed $0.9T_{\text{crit}}$. For these temperatures, coexistence densities are determined by Gaussian fits to density probability histograms, which are then used in fits to the scaling law (with a critical exponent of 0.326) and law of rectilinear diameters to obtain the critical temperature and density. An Antoine fit over this same temperature range is used for the estimation of the critical pressure. The normal boiling point is determined from a Clausius-Clapeyron fit to the vapor pressure data for those temperatures closest to the normal boiling point. Specific information is provided for each validation molecule in Table [S1](#).

Table S1: **Simulation Details for TraPPE Validation Compounds**

Information is provided for the year the validation was performed, the system size (N), the number of independent simulations (N_{IND}), the total number of production cycles (N_{cyc}), the number of temperatures above $0.9T_{\text{crit}}$ that are used to determine the critical properties ($N_{T_{\text{crit}}}$), and the number of temperatures near T_{b} that are used to determine the normal boiling point ($N_{T_{\text{b}}}$). When a range of values is given, the larger/longer simulations are those closest to the critical point.

Molecule	Year	N	N_{IND}	$N_{\text{cyc}}/10^3$	$N_{T_{\text{crit}}}$	$N_{T_{\text{b}}}$
oxygen	2015	1200	8	200	5	2
carbon dioxide	2015	1000	8	50	4	2
ethane ^a	2017	1500	8	1000	4	2
<i>n</i> -butane	2019	400-800	8	100-250	3	2
<i>n</i> -decane ^b	2015	300	8	200	4	2
2-methylpropane	2019	400	8	100-200	4	3
2,2-dimethylpropane	2015	500	8	60	4	2
2-methylpropene	2014	500	8	60	4	2
2-methyl-1,3-butadiene	2013	500	8	80	4	3
naphthalene	2013	400	8	100	5	3
ethanol	2019	1000	8	80	3	2
2-butanol	2019	1000	8	100	5	2
ethanal	2018	800	8	100-150	4	3
pentanal	2019	400-800	16	120-280	5	4
acetone	2013	500	8	120	6	3
2-methoxy-2-methylpropane	2014	500	8	80	4	2
methyl acetate	2019	500-600	16	200	5	4
2-butanethiol	2019	500	8	100	4	2
dimethyl sulfide	2018	800	8	100-150	4	3
<i>n</i> -perfluoropentane	2019	400	8-16	100-200	4	4

^a The validation simulations for ethane were previously reported in: Shah, M. S.; Tsapatis, M.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. Improved United-atom Description of Ethane and Ethylene. *AIChE J.* **2017**, *63*, 5098–5110.

^b The validation simulations for *n*-decane were previously reported in: Dinpajoo, M.; Bai, P.; Allan, D. A.; Siepmann, J. I. Accurate and Precise Determination of Critical Properties from Gibbs Ensemble Monte Carlo Simulations. *J. Chem. Phys.* **2015**, *143*, 114113.

VLCC Data for Validation Models

Table S2: **Vapor-Liquid Coexistence Data for TraPPE Validation Compounds**

When available, Z_i values computed from instantaneous z values including histogram analysis are also provided. Subscripts provide the uncertainties in the last digit(s), given as the standard error of the mean estimated from the set of independent simulations (i.e., 31.6₂ stands for 31.6 ± 0.2 and 5679₃₄ stands for 5679 ± 34)

molecule	T [K]	ρ_{liq} [g/ml]	ρ_{vap} [g/ml]	P_{vap} [kPa]	Z_i
oxygen	80	1.1854 ₁	0.00154 ₁	31.6 ₂	
	100	1.0826 ₁	0.01117 ₇	272 ₂	
	120	0.9639 ₁	0.0423 ₄	1101 ₉	
	139	0.8147 ₆	0.1169 ₅	2835 ₇	
	142	0.7841 ₉	0.136 ₂	3210 ₂₀	
	145	0.7536 ₆	0.165 ₂	3690 ₁₀	
	148	0.712 ₁	0.192 ₂	4120 ₁₀	
carbon dioxide	220	1.1586 ₄	0.0161 ₂	616 ₅	
	230	1.1219 ₅	0.0238 ₂	922 ₇	
	240	1.0836 ₄	0.0341 ₁	1330 ₅	
	250	1.0432 ₂	0.0485 ₄	1874 ₁₁	
	260	0.999 ₁	0.0673 ₂	2557 ₆	
	270	0.951 ₁	0.095 ₁	3447 ₁₄	
	275	0.923 ₁	0.109 ₁	3870 ₂₄	
	280	0.896 ₁	0.128 ₁	4412 ₂₄	
	285	0.866 ₁	0.157 ₂	5050 ₃₂	
	290	0.834 ₁	0.196 ₄	5679 ₃₄	
ethane	178	0.5511 ₁	0.00233 ₁	111.2 ₄	
	197	0.52643 ₃	0.00541 ₁	276.3 ₆	
	212	0.50573 ₃	0.00942 ₁	500.0 ₄	
	236	0.46940 ₃	0.02019 ₃	1102 ₂	
	256	0.43442 ₅	0.0354 ₁	1901 ₂	
	275	0.39405 ₄	0.0599 ₆	2995 ₁₅	
	280	0.3814 ₁	0.0683 ₄	3320 ₁₀	
	285	0.3676 ₁	0.079 ₁	3680 ₁₀	
	290	0.3524 ₃	0.093 ₁	4080 ₁₀	

Table S2: (continued)

molecule	T [K]	ρ_{liq} [g/ml]	ρ_{vap} [g/ml]	P_{vap} [kPa]	Z_i
<i>n</i> -butane	204	0.67206 ₁₇	0.0001884 ₄₂	5.47 ₁₃	0.99581 ₁₀
	234	0.64232 ₁₅	0.000938 ₁₄	30.90 ₄₄	0.98494 ₃₈
	264	0.61093 ₁₆	0.003050 ₃₀	110.9 ₁₀	0.96314 ₃₉
	294	0.57735 ₂₀	0.00774 ₁₀	300.3 ₂₉	0.92354 ₆₁
	324	0.54024 ₄₀	0.01668 ₂₄	668 ₁₀	0.8676 ₁₃
	343	0.51386 ₅₄	0.02577 ₃₂	1031 ₁₀	0.8189 ₁₁
	354	0.49690 ₃₇	0.03235 ₄₁	1287 ₁₀	0.7913 ₂₂
	384	0.44432 ₂₈	0.05768 ₁₄	2209 ₁₀	0.6998 ₁₃
	392	0.42652 ₄₁	0.0686 ₁₀	2544 ₁₅	0.6643 ₃₅
	400	0.40651 ₃₇	0.0814 ₁₁	2907 ₁₉	0.6284 ₄₇
	408	0.3841 ₁₁	0.1003 ₁₀	3318 ₁₄	0.5697 ₂₅
<i>n</i> -decane	400	0.65236 ₂₃	0.001610 ₂₀	36.6 ₅	
	420	0.63502 ₂₄	0.00281 ₆	66.3 ₁	
	440	0.6164 ₃	0.00459 ₇	111 ₂	
	460	0.5971 ₄	0.00717 ₁₃	177 ₃	
	480	0.5770 ₃	0.01090 ₁₅	273 ₃	
	500	0.5551 ₃	0.01605 ₂₂	402 ₄	
	520	0.5316 ₄	0.0230 ₄	571 ₇	
	540	0.5059 ₆	0.0329 ₆	798 ₁₀	
	560	0.4761 ₇	0.0462 ₁₁	1078 ₁₉	
	580	0.4411 ₇	0.0656 ₁₁	1451 ₁₆	
	600	0.3956 ₁₂	0.0946 ₁₄	1859 ₁₈	
2-methylpropane	198	0.65805 ₁₇	0.0002214 ₄₁	6.24 ₁₁	0.99558 ₁₀
	228	0.62736 ₅	0.001088 ₃	34.87 ₁₀	0.98328 ₁₂
	258	0.59519 ₃	0.0035549 ₅₄	125.60 ₁₉	0.95821 ₄₀
	278	0.57243 ₈	0.006767 ₃₁	250 ₁	0.93300 ₅₅
	308	0.53527 ₁₀	0.015166 ₅₀	584 ₂	0.87747 ₆₄
	323	0.51453 ₉	0.02148 ₈	832 ₃	0.84045 ₆₅
	338	0.49222 ₁₀	0.03005 ₁₁	1156 ₃	0.8001 ₁₁
	343	0.48495 ₁₈	0.03400 ₁₅	1298 ₄	0.7823 ₁₀
	353	0.46703 ₂₅	0.04154 ₁₄	1567 ₄	0.7519 ₁₀
	368	0.44129 ₄₀	0.05697 ₃₇	2077 ₉	0.6934 ₁₆
	376	0.42281 ₄₇	0.06599 ₇₃	2367 ₁₆	0.6689 ₃₃
	384	0.40401 ₅₂	0.0793 ₁₀	2721 ₁₈	0.6272 ₄₀
	390	0.38710 ₅₀	0.0933 ₁₆	3030 ₂₁	0.5864 ₅₅

Table S2: (continued)

molecule	T [K]	ρ_{liq} [g/ml]	ρ_{vap} [g/ml]	P_{vap} [kPa]	Z_i
2,2-dimethylpropane	270	0.6135 ₈	0.00324 ₂	96.9 ₄	
	300	0.5804 ₁	0.00787 ₅	252 ₂	
	330	0.5444 ₇	0.01668 ₅	551 ₁	
	360	0.5025 ₂	0.0321 ₂	1055 ₄	
	390	0.4501 ₄	0.0585 ₉	1820 ₁₆	
	400	0.4283 ₅	0.0710 ₇	2131 ₁₀	
	410	0.4050 ₈	0.091 ₂	2550 ₂₃	
	420	0.373 ₂	0.114 ₅	2930 ₃₀	
2-methylpropene	253	0.64146 ₈	0.00236 ₂	86.0 ₆	
	274	0.61759 ₈	0.00490 ₃	188.3 ₁₀	
	299	0.58749 ₁₄	0.01014 ₆	408 ₂	
	329	0.5468 ₃	0.0213 ₂	877 ₇	
	350	0.5141 ₃	0.0336 ₃	1369 ₈	
	375	0.4678 ₄	0.0570 ₈	2197 ₁₉	
	380	0.4572 ₄	0.0625 ₆	2384 ₁₆	
	385	0.4454 ₈	0.0711 ₈	2621 ₁₉	
	390	0.4328 ₉	0.0774 ₁₀	2831 ₁₈	
2-methyl-1,3-butadiene	275	0.6962 ₁	0.00135 ₁	44.4 ₂	
	300	0.67044 ₇	0.00318 ₁	112.0 ₄	
	325	0.6431 ₂	0.00645 ₃	239.4 ₉	
	400	0.5456 ₃	0.0344 ₂	1321 ₅	
	435	0.4831 ₅	0.0665 ₆	2380 ₁₀	
	445	0.4602 ₈	0.0801 ₈	2750 ₁₀	
	450	0.4464 ₈	0.089 ₁	2980 ₂₀	
	455	0.433 ₁	0.101 ₁	3220 ₂₀	
naphthalene	475	0.8895 ₂	0.00297 ₃	88.6 ₉	
	500	0.8673 ₂	0.00495 ₃	153 ₁	
	550	0.8202 ₂	0.01216 ₅	393 ₁	
	600	0.7679 ₄	0.0257 ₂	841 ₆	
	650	0.7058 ₆	0.0504 ₄	1600 ₉	
	680	0.6630 ₄	0.0762 ₉	2280 ₁₀	
	690	0.6477 ₆	0.090 ₁	2560 ₂₀	
	700	0.627 ₁	0.100 ₂	2800 ₃₀	
	710	0.614 ₁	0.116 ₃	3120 ₃₀	
	720	0.593 ₁	0.134 ₂	3440 ₁₀	

Table S2: (continued)

molecule	T [K]	ρ_{liq} [g/ml]	ρ_{vap} [g/ml]	P_{vap} [kPa]	Z_i
ethanol	300	0.781 ₂	0.000152 ₁	8.12 ₄	
	325	0.757 ₂	0.000546 ₃	30.8 ₂	
	350	0.732 ₃	0.00162 ₁	95.8 ₄	
	375	0.704 ₂	0.0041 ₁	245 ₆	
	400	0.672 ₄	0.0088 ₁	538 ₆	
	425	0.634 ₃	0.0178 ₄	1065 ₂₂	
	450	0.589 ₅	0.034 ₁	1934 ₃₂	
	475	0.527 ₇	0.061 ₂	3190 ₇₁	
	485	0.497 ₉	0.083 ₃	3906 ₇₅	
	495	0.454 ₁₆	0.108 ₄	4716 ₈₉	
2-butanol	375	0.72416 ₃₆	0.003142 ₃₄	124.9 ₁₅	0.9455 ₂₃
	400	0.69253 ₃₈	0.00724 ₁₀	294.2 ₃₇	0.9070 ₃₄
	425	0.65617 ₅₅	0.01489 ₁₈	605.0 ₆₂	0.8537 ₂₆
	450	0.61440 ₇₅	0.02837 ₃₈	1127 ₁₂	0.7895 ₂₈
	460	0.59458 ₇₅	0.03570 ₄₈	1396 ₁₁	0.7600 ₄₄
	470	0.57385 ₆₉	0.04446 ₁₂	1706 ₃₁	0.7271 ₆₄
	480	0.5506 ₁₄	0.0560 ₁₈	2080 ₃₄	0.6855 ₆₉
	490	0.5237 ₁₆	0.0690 ₁₄	2488 ₂₄	0.652 ₁₀
	500	0.4943 ₂₈	0.0915 ₄₀	3003 ₄₅	0.588 ₁₀
	510	0.4537 ₄₃	0.1154 ₅₉	3530 ₄₅	0.533 ₁₆
ethanal	220	0.8761 ₂	0.0000529 ₁₉	2.18 ₈	0.9962 ₁₀
	240	0.8523 ₃	0.000184 ₆	8.27 ₂₅	0.9914 ₁₃
	260	0.8275 ₃	0.000517 ₈	24.9 ₄	0.9826 ₁₁
	280	0.8023 ₄	0.001234 ₁₉	63.1 ₉	0.9681 ₁₈
	300	0.7760 ₄	0.00258 ₄	138.2 ₂₂	0.9474 ₁₀
	320	0.7486 ₄	0.00492 ₅	273.1 ₂₅	0.9205 ₁₈
	340	0.7197 ₅	0.00872 ₁₀	495 ₅	0.8849 ₁₃
	360	0.6880 ₃	0.01435 ₁₀	822 ₅	0.8446 ₁₅
	380	0.6537 ₁₀	0.0231 ₃	1310 ₁₅	0.7932 ₂₀
	400	0.6156 ₅	0.0360 ₇	1984 ₂₇	0.7332 ₄₂
	420	0.5721 ₁₇	0.0554 ₁₃	2900 ₄₀	0.658 ₇
	425	0.5585 ₁₇	0.0608 ₂₀	3160 ₅₀	0.645 ₁₀
	430	0.5459 ₂₀	0.0676 ₁₅	3430 ₄₀	0.623 ₈
	435	0.5307 ₁₄	0.0759 ₂₃	3740 ₅₀	0.600 ₁₁

Table S2: (continued)

molecule	T [K]	ρ_{liq} [g/ml]	ρ_{vap} [g/ml]	P_{vap} [kPa]	Z_i
pentanal	300	0.80526 ₈	0.0002305 ₁₄	6.63 ₄	0.99407 ₁₂
	330	0.77562 ₇	0.000819 ₃	25.67 ₁₁	0.98444 ₂₁
	360	0.74507 ₇	0.002303 ₉	77.4 ₃	0.96691 ₁₇
	390	0.71208 ₈	0.00530 ₂	187.4 ₅	0.93955 ₃₉
	420	0.67664 ₁₃	0.01084 ₅	396 ₂	0.89967 ₆₀
	450	0.6377 ₁₃	0.02007 ₁₁	740 ₃	0.8473 ₁₀
	480	0.5929 ₃	0.0354 ₂	1283 ₅	0.7830 ₁₃
	510	0.5391 ₄	0.0610 ₅	2077 ₁₀	0.6934 ₂₃
	520	0.5172 ₆	0.0735 ₅	2403 ₁₁	0.6539 ₂₀
	530	0.4922 ₅	0.0890 ₅	2768 ₇	0.6103 ₂₁
	540	0.4635 ₈	0.1084 ₁₃	3188 ₉	0.557 ₁₃
	545	0.4433 ₉	0.120 ₃	3400 ₂₀	0.545 ₁₀
acetone	300	0.7753 ₂	0.00135 ₁	56.6 ₄	
	340	0.7308 ₂	0.00471 ₃	217 ₁	
	380	0.6820 ₂	0.01250 ₄	605 ₂	
	420	0.6264 ₂	0.0285 ₁	1385 ₅	
	460	0.5572 ₆	0.0602 ₂	2740 ₂₀	
	470	0.5358 ₅	0.0718 ₆	3180 ₁₀	
	475	0.5258 ₇	0.081 ₁	3450 ₂₀	
	480	0.5107 ₉	0.0853 ₆	3640 ₁₀	
	485	0.4985 ₆	0.097 ₂	3940 ₃₀	
	490	0.483 ₂	0.107 ₃	4210 ₃₀	
2-methoxy-2-methylpropane	495	0.464 ₃	0.114 ₂	4460 ₃₀	
	300	0.7464 ₁	0.00213 ₃	58.8 ₇	
	325	0.7180 ₂	0.00464 ₇	136 ₂	
	350	0.6877 ₁	0.00919 ₄	279 ₁	
	400	0.6187 ₄	0.0273 ₂	857 ₅	
	450	0.5263 ₈	0.072 ₁	2050 ₂₀	
	460	0.5014 ₃	0.0857 ₆	2380 ₁₀	
	465	0.487 ₁	0.094 ₁	2540 ₂₀	
	470	0.475 ₂	0.110 ₃	2800 ₃₀	

Table S2: (continued)

molecule	T [K]	ρ_{liq} [g/ml]	ρ_{vap} [g/ml]	P_{vap} [kPa]	Z_i
methyl acetate	260	0.94896 ₁₂	0.000173 ₄	5.04 ₁₁	0.99657 ₁₁
	295	0.90907 ₁₃	0.000850 ₁₂	27.8 ₄	0.98817 ₂₄
	330	0.86706 ₁₀	0.00291 ₃	104.5 ₉	0.96932 ₃₁
	365	0.82244 ₁₀	0.00763 ₄	292.9 ₁₄	0.93642 ₂₅
	400	0.77326 ₁₀	0.01693 ₉	674 ₃	0.88648 ₄₆
	435	0.7177 ₃	0.03324 ₁₄	1334 ₄	0.8223 ₁₀
	470	0.6516 ₄	0.0630 ₃	2412 ₉	0.7279 ₁₄
	480	0.6297 ₄	0.0745 ₄	2794 ₁₀	0.6971 ₁₁
	490	0.6043 ₅	0.0888 ₆	3230 ₁₁	0.6629 ₂₀
	500	0.5769 ₉	0.1061 ₇	3701 ₁₂	0.6235 ₂₀
	510	0.5434 ₇	0.1290 ₇	4246 ₁₃	0.5788 ₂₄
2-butanethiol	280	0.84987 ₁₁	0.0002081 ₂₀	5.358 ₅₁	0.99616 ₁₂
	300	0.830242 ₈₅	0.000514 ₁₀	14.09 ₂₇	0.99211 ₁₉
	320	0.810012 ₇₂	0.0011012 ₄₇	32.00 ₁₃	0.98521 ₂₂
	340	0.789451 ₆₉	0.0021309 ₇₇	65.13 ₂₄	0.97534 ₂₄
	360	0.76835 ₁₁	0.003814 ₁₆	121.55 ₄₈	0.96079 ₂₄
	380	0.74624 ₁₅	0.006344 ₄₇	209.2 ₁₄	0.94266 ₆₁
	400	0.72309 ₁₂	0.010022 ₆₄	339.6 ₂₀	0.92042 ₅₄
	420	0.69859 ₁₀	0.015056 ₆₀	520.0 ₂₀	0.89382 ₈₁
	440	0.67308 ₁₈	0.022291 ₈₉	773.4 ₂₅	0.85808 ₈₃
	460	0.64455 ₁₈	0.03172 ₁₆	1094.0 ₄₁	0.8175 ₁₁
	480	0.61345 ₂₃	0.04457 ₁₀	1512.2 ₂₅	0.7723 ₁₀
	500	0.57837 ₂₉	0.06204 ₃₇	2030.3 ₇₀	0.7190 ₂₁
	510	0.55816 ₅₈	0.07331 ₅₃	2333 ₁₁	0.6912 ₄₇
	520	0.53510 ₉₃	0.08613 ₆₉	2662 ₁₂	0.6670 ₂₈
	525	0.52585 ₅₁	0.0981 ₁₁	2894 ₁₂	0.6373 ₆₀
	530	0.51110 ₇₈	0.1054 ₁₀	3055 ₁₀	0.6228 ₄₂

Table S2: (continued)

molecule	T [K]	ρ_{liq} [g/ml]	ρ_{vap} [g/ml]	P_{vap} [kPa]	Z_i
dimethyl sulfide	255	0.8856 ₄	0.000385 ₁₂	13.1 ₄	0.99415 ₃₀
	290	0.84597 ₉	0.00163 ₂	62.0 ₉	0.98123 ₆₂
	325	0.8042 ₃	0.00486 ₇	202 ₃	0.9551 ₁₃
	360	0.75948 ₁₄	0.01175 ₉	517 ₃	0.9135 ₁₆
	395	0.7099 ₂	0.02427 ₁₅	1094 ₇	0.8544 ₁₀
	430	0.6524 ₅	0.0463 ₄	2052 ₁₂	0.7734 ₃₁
	455	0.603 ₅	0.0711 ₇	3050 ₃₀	0.7036 ₅₈
	465	0.5801 ₇	0.0837 ₁₃	3505 ₂₆	0.6712 ₅₇
	475	0.5538 ₁₇	0.1004 ₂₃	4030 ₅₀	0.633 ₁₀
	485	0.525 ₃	0.124 ₆	4640 ₆₀	0.581 ₂₁
<i>n</i> -perfluoropentane	200	1.9183 ₇	0.000419 ₂₅	2.41 ₁₅	0.9966 ₄
	220	1.8625 ₉	0.00134 ₃	8.42 ₁₈	0.9913 ₅
	240	1.8048 ₆	0.00347 ₃	23.62 ₁₇	0.9811 ₅
	260	1.7450 ₉	0.00764 ₇	55.3 ₅	0.9658 ₉
	280	1.6827 ₁₂	0.01483 ₁₉	112.8 ₁₃	0.9424 ₁₇
	310	1.5805 ₁₃	0.0341 ₅	273 ₃	0.896 ₃
	340	1.4645 ₁₈	0.0697 ₈	563 ₅	0.827 ₃
	370	1.3243 ₁₃	0.1324 ₁₇	1025 ₁₁	0.734 ₄
	380	1.272 ₂	0.1602 ₁₆	1236 ₈	0.700 ₅
	385	1.241 ₂	0.178 ₄	1349 ₁₅	0.687 ₆
	390	1.208 ₄	0.199 ₄	1475 ₁₇	0.660 ₆
	395	1.173 ₆	0.224 ₆	1600 ₂₀	0.618 ₈

Table S3: **Critical Properties and Normal Boiling Points for TraPPE Validation Compounds**

Subscripts provide the uncertainties in the last digit(s), given as the standard error of the mean estimated from the set of independent simulations (i.e., 302.8₄ stands for 302.8 ± 0.4 and 0.2354₁₁ stands for 0.2354 ± 0.0011)

Molecule	T_{crit} [K]	ρ_{crit} [g/ml]	P_{crit} [MPa]	Z_{crit}	T_{boil}
oxygen	153.5 ₁	0.440 ₁	5.06 ₃	0.288 ₂	89.73 ₄
carbon dioxide	303.7 ₅	0.507 ₃	7.8 ₄	0.268 ₁₄	—
ethane ^a	302.8 ₄	0.218 ₁	5.2 ₁	0.285 ₆	176.2 ₁
<i>n</i> -butane	423.0 ₅	0.2354 ₁₁	4.20 ₁₀	0.295 ₇	262.0 ₂
<i>n</i> -decane ^b	619.8 ₇	0.237 ₂	2.38 ₃	0.277 ₅	436.3 ₁
2-methylpropane	407.3 ₆	0.2313 ₁₄	3.93 ₁₈	0.292 ₁₄	252.5 ₄
2,2-dimethylpropane	431 ₁	0.238 ₂	3.47 ₉	0.293 ₈	271.2 ₁
2-methylpropene	417 ₁	0.243 ₃	4.3 ₆	0.286 ₄₁	256.7 ₁
2-methyl-1,3-butadiene	475.1 ₈	0.258 ₂	4.2 ₂	0.281 ₁₄	297.3 ₇
naphthalene	755 ₁	0.357 ₂	4.7 ₂	0.269 ₁₂	480.9 ₃
ethanol	509 ₁	0.275 ₇	6.0 ₇	0.238 ₂₈	346 ₂
2-butanol	523.4 ₁₂	0.2767 ₃₂	4.40 ₁₆	0.271 ₁₀	369.7 ₃
ethanal	466 ₃	0.281 ₆	6 ₂	0.243 ₈₂	291.8 ₂
pentanal	559.8 ₄	0.2761 ₁₀	4.12 ₆	0.276 ₅	369.0 ₁
acetone	512.3 ₉	0.283 ₂	5.6 ₂	0.270 ₁₀	316.1 ₁
2-methoxy-2-methylpropane	491 ₁	0.278 ₃	3.7 ₂	0.287 ₁₆	315.8 ₃
methyl acetate	530.6 ₃	0.3253 ₉	5.49 ₇	0.283 ₄	328.8 ₂
2-butanethiol	557 ₂	0.299 ₄	4.3 ₆	0.280 ₄₀	353.9 ₂
dimethyl sulfide	507 ₂	0.313 ₅	6.0 ₄	0.283 ₁₉	303.6 ₃
<i>n</i> -perfluoropentane	419 ₂	0.67 ₁	2.3 ₄	0.284 ₅₀	277.0 ₂

Description of GitHub Repositories

The software packages used to generate and analyze the data in this manuscript are publicly available online as GitHub repositories.

Monte Carlo for Complex Chemical Systems–Minnesota (MCCCS–MN) is a Fortran software package developed by the Siepmann research group. The source code and `makefile` for the specific version of MCCCS–MN used for the validation simulations of pentanal, as well as sample input files (`fort.4`, `fort.77`, and `topmon.inp`) and output files (`run1a.dat` and `fort.12`) for a short production run of pentanal at 510 K in the *NVT*-Gibbs ensemble, are available at <https://github.com/SiepmannGroup/VLE-Validation>.

`VLE_Outlier` allows for the detection of outliers in VLE simulations using trends in the compressibility factor Z and Clausius-Clapeyron triads. To guard against typographical mistakes, `VLE_Outlier` compares the Z_{sim} value provided in the input file (for GEMC simulations, Z_{sim} should be calculated using Eqn. 5 and, only for $T_r > 0.9$, using histogram analysis) to that computed using Eqn. 4. Then, $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ (where Z_{exd} is taken from Eqn. 7) is compared to the prescribed bounds for the specific T_r range to determine whether Z_{sim} is an outlier. Because the Z metric (particularly, when Z_i from Eqn. 5 is used) can fail to detect inadequate sampling in GEMC simulations, `VLE_Outlier` also checks for outliers based on Clausius-Clapeyron triads. This is done by applying Eqn. 8 within triads of data, identifying P_{CC} and testing whether $\ln(P_{\text{sim}}/P_{\text{CC}})$ falls within the prescribed bounds. `VLE_Outlier` is available at <https://github.com/SiepmannGroup/VLE-Validation>.

`GEMC_Hist` is an analysis package for utilizing histograms to calculate accurate vapor and liquid coexistence densities, saturated vapor pressure, and compressibility factor from a Gibbs ensemble Monte Carlo trajectory for unary vapor–liquid equilibria near the critical point. `GEMC_Hist` is available at https://github.com/dejac001/GEMC_histogram_analysis with documentation available at https://dejac001.github.io/GEMC_histogram_analysis/. The exact version used of `GEMC_hist` used for making Figures 2-3 and Table 3 can be downloaded at https://github.com/dejac001/GEMC_histogram_analysis/archive/master.zip.

Outliers based on Compressibility Factor and CC Triads

Table S4: **All $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ Outliers in the TraPPE Development Data**

Outliers are defined as having $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ less than -0.03 or greater than 0.03 when $T_r \leq 0.70$, less than -0.06 or greater than 0.05 when $0.70 < T_r \leq 0.85$, and less than -0.09 or greater than 0.11 when $T_r > 0.85$. Uncertainties (unc) are provided as the standard error of the mean. The nearest boundary that each outlier exceeds is also provided, for easy reference, in the table in column NB.

Molecule	T [K]	T_r	Z	unc	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	unc	NB
n-octane (UA)	540	0.947	0.674	0.022	0.117	0.033	0.11
ethene (UA)	184	0.650	0.917	0.042	-0.037	0.046	-0.03
	225	0.795	0.730	0.031	-0.145	0.043	-0.06
propene (UA)	182	0.503	0.967		-0.032		-0.03
1,5-hexadiene (UA)	324	0.653	0.980		0.031		0.03
ethylbenzene (UA)	397	0.640	0.928		-0.030		-0.03
	447	0.721	1.085		0.179		0.05
	497	0.802	0.935		0.111		0.05
propylbenzene (UA)	450	0.696	0.866		-0.066		-0.03
isopropylbenzene (UA)	450	0.700	0.958		0.038		0.03
o-xylene (UA)	400	0.632	1.101		0.136		0.03
	450	0.711	0.967		0.055		0.05
	595	0.940	0.721		0.154		0.11
m-xylene (UA)	451	0.719	0.963		0.057		0.05
	547	0.872	0.683		-0.091		-0.09
p-xylene (UA)	397	0.643	0.816		-0.157		-0.03
methanol (UA)	425	0.847	0.678	0.039	-0.145	0.057	-0.06
	475	0.946	0.366	0.017	-0.499	0.046	-0.09
propan-1-ol (UA)	500	0.929	0.570	0.065	-0.122	0.114	-0.09
pentan-1-ol (UA)	400	0.691	0.965	0.057	0.039	0.059	0.03
	450	0.777	0.922	0.078	0.068	0.084	0.05

Table S4: (continued)

Molecule	T [K]	T_r	Z	unc	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	unc	NB
octan-1-ol (UA)	300	0.477	1.160	0.136	0.149	0.117	0.03
	350	0.556	1.053	0.228	0.064	0.216	0.03
	450	0.715	0.963	0.073	0.055	0.076	0.05
	500	0.795	0.908	0.039	0.073	0.043	0.05
	600	0.954	0.680	0.043	0.157	0.064	0.11
propan-2-ol (UA)	300	0.598	0.688	0.099	-0.348	0.144	-0.03
	400	0.797	0.783	0.089	-0.072	0.113	-0.06
butan-2-ol (UA)	350	0.668	0.996	0.102	0.055	0.102	0.03
2-methylpropan-2-ol (UA)	350	0.694	0.959	0.093	0.035	0.097	0.03
	480	0.952	0.441	0.025	-0.284	0.057	-0.09
1,3-propanediol (UA)	450	0.620	1.001	0.078	0.036	0.078	0.03
	500	0.689	0.962	0.034	0.034	0.035	0.03
	550	0.758	0.924	0.076	0.051	0.082	0.05
pentane-1,5-diol (UA)	550	0.775	0.934	0.084	0.079	0.090	0.05
	600	0.845	0.854	0.043	0.083	0.051	0.05
dimethyl ether (UA)	329	0.827	0.755	0.021	-0.068	0.028	-0.06
ethanal (UA)	260	0.561	0.685		-0.364		-0.03
	300	0.647	0.893		-0.065		-0.03
	440	0.950	0.529		-0.114		-0.09
pentanal (UA)	380	0.672	0.983		0.044		0.03
octanal (UA)	600	0.957	0.511		-0.114		-0.09
acetone (UA)	300	0.591	0.895		-0.088		-0.03
	340	0.669	0.858		-0.093		-0.03
	380	0.748	0.830		-0.067		-0.06
2-octanone (UA)	360	0.562	0.575		-0.540		-0.03
	400	0.625	0.861		-0.113		-0.03
	520	0.812	0.883		0.068		0.05
acetonitrile (UA)	398	0.725	0.849	0.050	-0.064	0.059	-0.06
	448	0.816	0.730	0.031	-0.117	0.042	-0.06
	498	0.907	0.585	0.031	-0.163	0.053	-0.09

Table S4: (continued)

Molecule	T [K]	T_r	Z	unc	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	unc	NB
propionitrile (UA)	448	0.806	0.776	0.030	-0.070	0.039	-0.06
	498	0.896	0.625	0.034	-0.128	0.055	-0.09
pyrimidine (UA)	298	0.489	0.771	0.016	-0.260	0.021	-0.03
	365	0.598	1.075	0.051	0.098	0.048	0.03
	398	0.652	0.885	0.097	-0.072	0.109	-0.03
	480	0.787	0.588	0.128	-0.371	0.218	-0.06
nitrobenzene (UA)	298	0.399	0.022	0.108	-3.835	5.002	-0.03
	375	0.502	0.585	0.182	-0.534	0.311	-0.03
methanethiol (UA)	220	0.463	0.658		-0.419		-0.03
	240	0.505	0.891		-0.113		-0.03
	260	0.547	0.931		-0.062		-0.03
ethanethiol (UA)	240	0.478	1.112		0.106		0.03
pentanethiol (UA)	300	0.498	3.214		1.169		0.03
	320	0.532	1.843		0.618		0.03
	340	0.565	1.395		0.347		0.03
	360	0.598	1.192		0.202		0.03
	380	0.631	1.099		0.135		0.03
	400	0.664	1.035		0.092		0.03
	420	0.698	0.998		0.077		0.03
	440	0.731	0.958		0.063		0.05
octanethiol (UA)	460	0.764	0.921		0.054		0.05
	360	0.537	0.869		-0.133		-0.03
	380	0.566	0.950		-0.037		-0.03
2-butanethiol (UA)	460	0.686	0.961		0.031		0.03
	260	0.467	0.521		-0.651		-0.03
	280	0.503	0.704		-0.349		-0.03
	300	0.539	0.887		-0.112		-0.03
2-methyl-1-propanethiol (UA)	320	0.575	0.951		-0.032		-0.03
	300	0.539	0.443		-0.807		-0.03
	320	0.575	0.729		-0.299		-0.03
	340	0.610	0.855		-0.126		-0.03
	360	0.646	0.897		-0.062		-0.03

Table S4: (continued)

Molecule	T [K]	T_r	Z	unc	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	unc	NB
2-methyl-2-propanethiol (UA)	300	0.577	0.912		−0.074		−0.03
	320	0.615	0.920		−0.050		−0.03
	340	0.654	0.912		−0.041		−0.03
	360	0.692	0.897		−0.033		−0.03
dimethyl sulfide (UA)	240	0.476	0.934		−0.068		−0.03
	260	0.516	0.958		−0.039		−0.03
ethylmethyl sulfide (UA)	300	0.560	0.944		−0.044		−0.03
dimethyl disulfide (UA)	280	0.462	0.450		−0.799		−0.03
	300	0.495	0.657		−0.419		−0.03
	320	0.528	0.885		−0.116		−0.03
	340	0.561	0.929		−0.060		−0.03
diethyl disulfide (UA)	400	0.608	0.893		−0.084		−0.03
	420	0.638	0.919		−0.041		−0.03
thiophene (UA)	293	0.484	0.953		−0.048		−0.03
methyl acrylate (UA)	450	0.800	0.917	0.041	0.090	0.045	0.05
n-octyl acrylate (UA)	490	0.697	0.992	0.035	0.071	0.035	0.03
2-ethylhexyl acrylate (UA)	550	0.795	0.905	0.046	0.070	0.051	0.05
2-hydroxyethyl acrylate (UA)	470	0.693	0.958	0.024	0.033	0.025	0.03
n-butyl methacrylate (UA)	400	0.625	1.006	0.049	0.043	0.048	0.03
1,4-dioxane (UA)	555	0.944	0.539	0.035	−0.121	0.065	−0.09
SPC-pol3 (pol)	573	0.924	0.584	0.040	−0.115	0.068	−0.09
TIP4P-pol1 (pol)	373	0.643	1.005	0.126	0.051	0.125	0.03
	523	0.902	0.570	0.076	−0.205	0.134	−0.09
TIP4P-pol3 (pol)	523	0.902	0.614	0.036	−0.129	0.059	−0.09
n-octane (EH)	550	0.958	0.634	0.091	0.110	0.143	0.11
n-dodecane (EH)	450	0.675	0.996	0.084	0.060	0.085	0.03
	500	0.750	0.948	0.078	0.068	0.083	0.05
	550	0.825	0.902	0.054	0.106	0.060	0.05

Table S4: (continued)

Molecule	T [K]	T_r	Z	unc	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	unc	NB
methylamine (EH)	298	0.695	1.008	0.120	0.085	0.119	0.03
	350	0.816	0.696	0.012	-0.165	0.017	-0.06
dimethylamine (EH)	250	0.564	1.084	0.553	0.095	0.510	0.03
	280	0.632	1.007	1.009	0.047	1.002	0.03
	308	0.695	0.867	0.104	-0.066	0.120	-0.03
ethylamine (EH)	293	0.648	1.002	0.178	0.050	0.178	0.03
nitromethane (EH)	375	0.637	0.923	0.135	-0.038	0.146	-0.03
	550	0.934	0.575	0.030	-0.097	0.052	-0.09
nitroethane (EH)	323	0.555	1.048	0.175	0.059	0.167	0.03
acetonitrile (EH)	298	0.545	0.954	0.088	-0.038	0.092	-0.03
	398	0.728	0.843	0.021	-0.068	0.025	-0.06
	448	0.819	0.745	0.024	-0.094	0.032	-0.06
	498	0.910	0.529	0.033	-0.255	0.062	-0.09
propionitrile (EH)	498	0.897	0.631	0.036	-0.114	0.058	-0.09
benzene (EH)	315	0.560	1.023	0.417	0.035	0.408	0.03
	465	0.826	0.722	0.121	-0.115	0.168	-0.06
pyridine (EH)	298	0.482	0.912	0.600	-0.092	0.659	-0.03
pyrimidine (EH)	298	0.472	0.969	1.616	-0.031	1.667	-0.03
	575	0.910	0.615	0.061	-0.106	0.099	-0.09
pyrazine (EH)	348	0.565	1.107	0.313	0.116	0.283	0.03
pyridazine (EH)	298	0.392	1.616	1.035	0.480	0.640	0.03
	348	0.457	0.791	0.357	-0.235	0.452	-0.03
	398	0.523	0.864	0.169	-0.141	0.195	-0.03
	468	0.615	1.003	0.180	0.036	0.179	0.03
thiophene (EH)	293	0.499	0.967	0.207	-0.032	0.214	-0.03
	345	0.588	1.010	0.160	0.032	0.158	0.03
furan (EH)	250	0.506	1.048	0.182	0.049	0.173	0.03
pyrrole (EH)	300	0.461	0.690	0.164	-0.372	0.238	-0.03
	370	0.568	0.945	0.321	-0.041	0.339	-0.03
	520	0.799	0.776	0.116	-0.079	0.149	-0.06

Table S4: (continued)

Molecule	T [K]	T_r	Z	unc	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	unc	NB
thiazole (EH)	350	0.550	1.024	0.194	0.034	0.190	0.03
	400	0.629	1.035	0.213	0.073	0.206	0.03
	450	0.708	0.980	0.116	0.066	0.118	0.05
oxazole (EH)	300	0.557	0.850	0.199	-0.150	0.234	-0.03
imidazole (EH)	400	0.489	1.159	0.192	0.148	0.166	0.03
	450	0.550	1.031	0.546	0.041	0.530	0.03
pyrazole (EH)	350	0.475	1.097	0.113	0.092	0.103	0.03
	400	0.543	0.955	0.259	-0.037	0.271	-0.03
chlorobenzene (EH)	575	0.909	0.599	0.117	-0.136	0.195	-0.09
1,2-dichlorobenzene (EH)	398	0.570	0.937	0.143	-0.049	0.153	-0.03
1,3-dichlorobenzene (EH)	650	0.938	0.562	0.059	-0.103	0.105	-0.09
1,2,3-trichlorobenzene (EH)	400	0.532	0.818	0.682	-0.194	0.833	-0.03
	700	0.931	0.582	0.124	-0.095	0.212	-0.09
	725	0.964	0.394	0.096	-0.331	0.243	-0.09
hexachlorobenzene (EH)	500	0.564	1.042	0.435	0.056	0.417	0.03
	575	0.648	0.894	0.372	-0.064	0.417	-0.03
1,2-dihydroxybenzene (EH)	500	0.619	1.059	0.276	0.092	0.260	0.03
	520	0.644	1.019	0.326	0.065	0.320	0.03
1,3-dihydroxybenzene (EH)	520	0.625	1.007	0.142	0.044	0.141	0.03
1,4-dihydroxybenzene (EH)	580	0.683	0.979	0.286	0.048	0.293	0.03
benzonitrile (EH)	298	0.435	1.189	0.180	0.173	0.151	0.03
	540	0.787	0.738	0.144	-0.143	0.195	-0.06
	640	0.933	0.554	0.086	-0.138	0.155	-0.09
p-benzoquinone (EH)	395	0.568	0.944	0.093	-0.043	0.098	-0.03
	430	0.618	1.002	0.073	0.035	0.073	0.03
	460	0.661	0.989	0.188	0.045	0.190	0.03
naphthalene (EH)	700	0.933	0.550	0.059	-0.142	0.107	-0.09

Table S4: (continued)

Molecule	T [K]	T_r	Z	unc	$\ln (Z_{\text{sim}}/Z_{\text{exd}})$	unc	NB
anthracene (EH)	550	0.615	0.612	0.172	-0.458	0.281	-0.03
	600	0.670	0.590	0.119	-0.466	0.201	-0.03
	625	0.698	0.598	0.113	-0.435	0.188	-0.03
	650	0.726	0.587	0.078	-0.432	0.133	-0.06
	700	0.782	0.544	0.128	-0.453	0.236	-0.06
	750	0.838	0.485	0.080	-0.493	0.166	-0.06
	800	0.894	0.414	0.135	-0.544	0.325	-0.09
phenanthrene (EH)	550	0.613	0.607	0.040	-0.467	0.067	-0.03
	600	0.669	0.603	0.119	-0.445	0.197	-0.03
	625	0.697	0.564	0.066	-0.495	0.117	-0.03
	650	0.725	0.559	0.081	-0.481	0.144	-0.06
	700	0.780	0.549	0.094	-0.447	0.171	-0.06
	750	0.836	0.500	0.055	-0.466	0.110	-0.06
	800	0.892	0.433	0.074	-0.503	0.170	-0.09
naphthalen-2-ol (EH)	825	0.920	0.405	0.037	-0.496	0.091	-0.09
	500	0.588	0.650	0.173	-0.408	0.266	-0.03
	550	0.647	0.631	0.167	-0.413	0.265	-0.03
	575	0.676	0.635	0.207	-0.389	0.327	-0.03
	600	0.706	0.635	0.127	-0.369	0.200	-0.06
	650	0.765	0.576	0.045	-0.414	0.078	-0.06
	700	0.824	0.519	0.115	-0.448	0.221	-0.06
naphthalene-2-carbonitrile (EH)	750	0.882	0.468	0.100	-0.449	0.214	-0.09
	775	0.912	0.430	0.084	-0.458	0.195	-0.09
	450	0.534	0.682	0.089	-0.375	0.130	-0.03
	500	0.593	0.676	0.167	-0.369	0.247	-0.03
	525	0.623	0.643	0.217	-0.405	0.338	-0.03
	550	0.652	0.609	0.206	-0.445	0.338	-0.03
	580	0.688	0.665	0.107	-0.336	0.161	-0.03
quinoline (EH)	600	0.712	0.594	0.076	-0.431	0.128	-0.06
	650	0.771	0.582	0.078	-0.398	0.134	-0.06
	700	0.830	0.550	0.064	-0.380	0.116	-0.06
	750	0.890	0.441	0.101	-0.491	0.229	-0.09
	300	0.373	0.964	0.022	-0.037	0.022	-0.03

Table S4: (continued)

Molecule	T [K]	T_r	Z	unc	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	unc	NB
indole (EH)	400	0.522	1.123	0.165	0.121	0.147	0.03
	450	0.587	1.060	0.184	0.079	0.174	0.03
	500	0.652	1.073	0.158	0.121	0.147	0.03
	550	0.717	1.032	0.193	0.125	0.187	0.05
	600	0.782	0.958	0.081	0.113	0.085	0.05
isoindole (EH)	400	0.485	0.896	0.150	-0.110	0.168	-0.03
	450	0.546	0.895	0.211	-0.102	0.236	-0.03
	500	0.607	0.916	0.208	-0.059	0.227	-0.03
	550	0.667	1.067	0.114	0.125	0.107	0.03
	600	0.728	0.842	0.187	-0.069	0.223	-0.06
	650	0.789	0.795	0.098	-0.067	0.123	-0.06
	700	0.850	0.707	0.062	-0.098	0.088	-0.06
	750	0.910	0.582	0.062	-0.161	0.107	-0.09
benzimidazole (EH)	500	0.531	1.089	0.952	0.092	0.874	0.03
	700	0.744	0.956	0.167	0.072	0.175	0.05
	850	0.903	0.635	0.122	-0.093	0.193	-0.09
indazole (EH)	500	0.579	1.089	0.492	0.105	0.451	0.03
	550	0.637	1.158	0.239	0.190	0.207	0.03
	800	0.927	0.582	0.090	-0.110	0.156	-0.09
purine (EH)	500	0.525	1.051	0.389	0.054	0.370	0.03
	550	0.578	1.103	0.574	0.117	0.521	0.03
	600	0.630	1.115	0.241	0.148	0.216	0.03
	875	0.919	0.607	0.070	-0.092	0.115	-0.09
benzo[c]thiophene (EH)	690	0.899	0.633	0.099	-0.107	0.156	-0.09
benzisoazole (EH)	600	0.733	1.075	0.161	0.179	0.149	0.05
	750	0.917	0.573	0.138	-0.157	0.240	-0.09
benzothiazole (EH)	450	0.572	1.023	0.256	0.039	0.251	0.03
	550	0.699	1.079	0.130	0.156	0.120	0.03

Table S5: **All $\ln(P_{\text{sim}}/P_{\text{CC}})$ Outliers in the TraPPE Development Data**

CC_{LOW} outliers are defined as having $\ln(P_{\text{sim}}/P_{\text{CC}})$ greater than 0.06 or less than -0.09 when $T_r \leq 0.70$, 0.05 and -0.06 when $0.70 < T_r \leq 0.85$, and 0.06 and -0.06 when $T_r > 0.85$. CC_{MID} outliers are defined as having $\ln(P_{\text{sim}}/P_{\text{CC}})$ greater than 0.03 or less than -0.01 when $T_r \leq 0.70$, 0.03 and -0.02 when $0.70 < T_r \leq 0.85$, and 0.03 and -0.03 when $T_r > 0.85$. CC_{HIGH} outliers are defined as having $\ln(P_{\text{sim}}/P_{\text{CC}})$ greater than 0.05 or less than -0.09 when $T_r \leq 0.70$, 0.03 and -0.05 when $0.70 < T_r \leq 0.85$, and 0.04 and -0.04 when $T_r > 0.85$. The extrapolation range, ER, is provided for CC_{LOW} and CC_{HIGH} as the ratio of two differences, in 1000/K: the difference between the predicted temperature and its nearest temperature divided by the difference of the two temperatures used to make the prediction. CC_{MID} involves interpolation. The nearest boundary that each outlier exceeds is also provided, for easy reference, in the table in column NB.

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
ethane (UA)	178	0.586	105 ₃	CC_{LOW}	-0.264	1.158	-0.09
	197	0.648	286 ₂₅	CC_{LOW}	0.248	1.261	0.06
				CC_{MID}	0.122	—	-0.01
	217	0.714	541 ₁₆	CC_{LOW}	-0.180	1.121	-0.06
				CC_{MID}	-0.110	—	-0.02
				CC_{HIGH}	-0.228	0.863	-0.05
	236	0.776	1092 ₇₅	CC_{LOW}	0.166	1.227	0.05
				CC_{MID}	0.085	—	-0.02
				CC_{HIGH}	0.197	0.793	-0.05
	256	0.842	1740 ₁₈₁	CC_{MID}	-0.075	—	-0.02
				CC_{HIGH}	-0.161	0.892	-0.05
	275	0.905	2913 ₂₅₄	CC_{HIGH}	0.135	0.815	0.04
propane (UA)	217	0.590	79 ₆	CC_{LOW}	0.124	1.295	0.06
	249	0.677	276 ₁₁	CC_{LOW}	-0.229	1.293	-0.09
				CC_{MID}	-0.054	—	-0.01
	281	0.764	798 ₄₂	CC_{LOW}	0.133	1.186	0.05
				CC_{MID}	0.100	—	-0.02
				CC_{HIGH}	0.096	0.772	-0.05
	312	0.848	1519 ₇₈	CC_{MID}	-0.061	—	-0.02
				CC_{HIGH}	-0.177	0.773	-0.05
	344	0.935	2925 ₁₀₀	CC_{HIGH}	0.112	0.843	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
n-butane (UA)	262	0.619	109 ₂₁	CC _{LOW}	0.291	1.287	0.06
	295	0.697	285 ₂₈	CC _{LOW}	-0.176	1.183	-0.09
				CC _{MID}	-0.127	—	-0.01
	327	0.773	754 ₁₀₁	CC _{MID}	0.081	—	0.03
				CC _{HIGH}	0.226	0.777	-0.05
	360	0.851	1478 ₁₇₇	CC _{HIGH}	-0.149	0.845	-0.04
n-pentane (UA)	298	0.634	105 ₄	CC _{LOW}	-0.208	1.318	-0.09
	336	0.715	338 ₁₅	CC _{LOW}	0.271	1.436	0.05
				CC _{MID}	0.090	—	-0.02
	372	0.791	701 ₃₈	CC _{LOW}	-0.224	0.957	-0.06
				CC _{MID}	-0.111	—	-0.02
				CC _{HIGH}	-0.158	0.759	-0.05
	402	0.855	1407 ₃₅	CC _{MID}	0.114	—	0.03
				CC _{HIGH}	0.189	0.697	-0.04
n-octane (UA)	439	0.934	2306 ₅₈	CC _{HIGH}	-0.234	1.045	-0.04
	440	0.772	353 ₇	CC _{LOW}	-0.160	1.227	-0.06
	490	0.860	908 ₇	CC _{MID}	0.072	—	0.03
n-dodecane (UA)	540	0.947	1721 ₂₁	CC _{HIGH}	-0.130	0.815	-0.04
	450	0.675	46 ₇	CC _{LOW}	-0.212	1.222	-0.09
	500	0.750	169 ₁₃	CC _{LOW}	0.132	1.671	0.05
				CC _{MID}	0.095	—	-0.02
	550	0.825	412 ₂₀	CC _{LOW}	-0.108	1.127	-0.06
				CC _{MID}	-0.050	—	-0.02
				CC _{HIGH}	-0.174	0.818	-0.05
	585	0.877	760 ₃₄	CC _{MID}	0.051	—	0.03
				CC _{HIGH}	0.079	0.598	-0.04
n-dodecane (UA)	620	0.930	1189 ₄₆	CC _{HIGH}	-0.096	0.887	-0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
2-methylpropane (UA)	232	0.567	45 ₁	CC _{LOW}	0.372	1.549	0.06
	271	0.663	176 ₄	CC _{LOW}	−0.110	1.546	−0.09
				CC _{MID}	−0.146	—	−0.01
	304	0.743	540 ₂₂	CC _{MID}	0.043	—	0.03
				CC _{HIGH}	0.240	0.646	−0.05
	330	0.807	1039 ₂₉	CC _{HIGH}	−0.071	0.647	−0.05
2,2-dimethylpropane (UA)	281	0.650	138 ₂	CC _{LOW}	−0.150	0.613	−0.09
	309	0.715	341 ₇	CC _{LOW}	0.206	2.753	0.05
				CC _{MID}	0.093	—	−0.02
	369	0.854	1169 ₃₀	CC _{MID}	−0.055	—	−0.03
				CC _{HIGH}	−0.244	1.632	−0.04
	397	0.918	1971 ₃₈	CC _{HIGH}	0.075	0.363	0.04
2,3-dimethylbutane (UA)	330	0.656	134 ₃	CC _{LOW}	0.322	1.592	0.06
	384	0.763	487 ₃₂	CC _{LOW}	−0.434	1.229	−0.06
				CC _{MID}	−0.124	—	−0.02
	428	0.851	1341 ₄₀	CC _{MID}	0.195	—	0.03
				CC _{HIGH}	0.202	0.628	−0.04
	472	0.938	2148 ₁₁₉	CC _{HIGH}	−0.353	0.814	−0.04
2,2-dimethylhexane (UA)	354	0.642	64 ₁	CC _{LOW}	−0.230	1.467	−0.09
	403	0.731	256 ₈	CC _{LOW}	0.077	1.051	0.05
				CC _{MID}	0.093	—	−0.02
	445	0.808	563 ₁₇	CC _{MID}	−0.037	—	−0.02
				CC _{HIGH}	−0.157	0.682	−0.05
	494	0.897	1282 ₄₇	CC _{HIGH}	0.073	0.952	0.04
2,5-dimethylhexane (UA)	349	0.624	47.1 ₇	CC _{LOW}	−0.127	4.085	−0.09
	440	0.787	469 ₁₆	CC _{LOW}	−0.063	1.136	−0.06
	470	0.841	798 ₁₈	CC _{LOW}	0.070	1.451	0.05
	500	0.894	1205 ₆₅	CC _{HIGH}	−0.056	0.880	−0.04
	523	0.936	1680 ₆₀	CC _{HIGH}	0.048	0.689	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
3,4-dimethylhexane (UA)	352	0.603	41.7 ₉	CC _{LOW}	−0.560	3.917	−0.09
	440	0.753	409 ₂₄	CC _{MID}	0.114	—	0.03
	470	0.805	635 ₂₇	CC _{LOW}	0.150	1.061	0.05
				CC _{MID}	−0.021	—	−0.02
				CC _{HIGH}	−0.143	0.255	−0.05
	500	0.856	973 ₁₇	CC _{MID}	−0.073	—	−0.03
	532	0.911	1675 ₇₂	CC _{HIGH}	0.141	0.942	0.04
ethene (UA)	144	0.509	27 ₁	CC _{LOW}	0.123	1.278	0.06
	164	0.580	95 ₃	CC _{MID}	−0.054	—	−0.01
	184	0.650	280 ₈	CC _{LOW}	−0.109	1.284	−0.09
				CC _{HIGH}	0.096	0.783	−0.09
	205	0.724	690 ₂₁	CC _{LOW}	0.128	1.195	0.05
				CC _{MID}	0.048	—	−0.02
	225	0.795	1280 ₃₈	CC _{MID}	−0.058	—	−0.02
				CC _{HIGH}	−0.085	0.779	−0.05
	245	0.866	2390 ₇₁	CC _{HIGH}	0.107	0.837	0.04
propene (UA)	245	0.677	310	CC _{LOW}	0.200	1.218	0.06
	276	0.762	760	CC _{MID}	−0.090	—	−0.02
	308	0.851	1870	CC _{HIGH}	0.164	0.821	0.04
1-butene (UA)	245	0.592	64 ₂	CC _{LOW}	−0.153	1.256	−0.09
	271	0.655	190 ₆	CC _{MID}	0.068	—	0.03
	296	0.715	400 ₁₂	CC _{HIGH}	−0.122	0.796	−0.05
	322	0.778	770 ₂₃	CC _{LOW}	0.079	1.114	0.05
	347	0.838	1320 ₄₀	CC _{MID}	−0.038	—	−0.02
	373	0.901	2300 ₆₉	CC _{HIGH}	0.071	0.898	0.04
cis-2-butene (UA)	245	0.563	47	CC _{LOW}	−0.218	2.434	−0.09
	295	0.678	290	CC _{LOW}	0.138	0.669	0.06
				CC _{MID}	0.064	—	−0.01
	322	0.740	560	CC _{MID}	−0.083	—	−0.02
				CC _{HIGH}	−0.090	0.411	−0.05
	373	0.857	1840	CC _{HIGH}	0.207	1.494	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
trans-2-butene (UA)	247	0.580	58	CC _{LOW}	0.065	0.641	0.06
	272	0.638	145	CC _{MID}	−0.040	—	−0.01
	323	0.758	670	CC _{HIGH}	0.101	1.560	0.03
2-methylpropene (UA)	274	0.656	190	CC _{MID}	0.045	—	0.03
	299	0.715	410	CC _{LOW}	0.153	1.672	0.05
				CC _{MID}	0.038	—	−0.02
				CC _{HIGH}	−0.090	1.007	−0.05
	329	0.787	820	CC _{MID}	−0.057	—	−0.02
				CC _{HIGH}	−0.076	0.999	−0.05
	350	0.837	1360	CC _{HIGH}	0.091	0.598	0.03
1-octene (UA)	393	0.693	130	CC _{LOW}	0.102	1.138	0.06
	441	0.778	390	CC _{MID}	−0.048	—	−0.02
	494	0.871	1120	CC _{HIGH}	0.090	0.878	0.04
1,5-hexadiene (UA)	289	0.583	21	CC _{LOW}	−0.325	1.630	−0.09
	324	0.653	90	CC _{LOW}	0.141	0.642	0.06
				CC _{MID}	0.124	—	−0.01
	350	0.706	180	CC _{LOW}	0.050	1.286	0.05
				CC _{MID}	−0.086	—	−0.02
				CC _{HIGH}	−0.200	0.613	−0.05
	400	0.806	660	CC _{MID}	−0.022	—	−0.02
				CC _{HIGH}	0.220	1.558	−0.05
benzene 6-site (UA)	300	0.531	25	CC _{LOW}	0.093	1.333	0.06
	350	0.619	130	CC _{LOW}	−0.205	1.286	−0.09
				CC _{MID}	−0.040	—	−0.01
	400	0.708	480	CC _{LOW}	0.085	1.250	0.05
				CC _{MID}	0.090	—	−0.02
				CC _{HIGH}	0.070	0.750	−0.05
	450	0.796	1130	CC _{MID}	−0.038	—	−0.02
				CC _{HIGH}	−0.160	0.778	−0.05
	500	0.885	2400	CC _{HIGH}	0.068	0.800	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
toluene 7-site (UA)	400	0.673	220	CC _{LOW}	0.108	1.250	0.06
	450	0.758	600	CC _{LOW}	-0.148	1.963	-0.06
				CC _{MID}	-0.048	—	-0.02
	500	0.842	1460	CC _{MID}	0.050	—	0.03
				CC _{HIGH}	0.087	0.800	-0.05
	530	0.892	2130	CC _{HIGH}	-0.075	0.509	-0.04
ethylbenzene (UA)	397	0.640	101	CC _{LOW}	-0.288	1.252	-0.09
	447	0.721	380	CC _{MID}	0.128	—	0.03
	497	0.802	870	CC _{HIGH}	-0.230	0.799	-0.05
propylbenzene (UA)	350	0.541	10	CC _{LOW}	0.072	1.286	0.06
	400	0.618	53	CC _{LOW}	-0.142	1.250	-0.09
				CC _{MID}	-0.031	—	-0.01
	450	0.696	205	CC _{MID}	0.063	—	0.03
				CC _{HIGH}	0.056	0.778	-0.09
	500	0.773	540	CC _{HIGH}	-0.114	0.800	-0.05
isopropylbenzene (UA)	400	0.622	59	CC _{LOW}	0.444	1.250	0.06
	450	0.700	164	CC _{LOW}	-0.132	1.222	-0.09
				CC _{MID}	-0.197	—	-0.01
	500	0.778	530	CC _{MID}	0.059	—	0.03
				CC _{HIGH}	0.355	0.800	-0.05
	550	0.855	1242	CC _{HIGH}	-0.108	0.818	-0.04
o-xylene (UA)	450	0.711	310	CC _{LOW}	0.065	1.222	0.05
	500	0.790	740	CC _{MID}	-0.029	—	-0.02
	550	0.869	1590	CC _{HIGH}	0.053	0.818	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
m-xylene (UA)	451	0.719	340	CC _{LOW}	−0.149	1.433	−0.06
				CC _{MID}	0.033	—	−0.02
	503	0.802	900	CC _{LOW}	0.091	1.455	0.05
				CC _{MID}	0.061	—	−0.02
				CC _{HIGH}	−0.058	0.776	−0.05
	547	0.872	1600	CC _{MID}	−0.037	—	−0.03
				CC _{HIGH}	−0.104	0.698	−0.04
	582	0.928	2530	CC _{HIGH}	0.063	0.687	0.04
p-xylene (UA)	397	0.643	104	CC _{LOW}	−0.135	1.471	−0.09
	452	0.733	370	CC _{LOW}	0.062	1.065	0.05
				CC _{MID}	0.055	—	−0.02
	499	0.809	800	CC _{MID}	−0.030	—	−0.02
				CC _{HIGH}	−0.092	0.680	−0.05
	553	0.896	1750	CC _{HIGH}	0.059	0.939	0.04
naphthalene (UA)	550	0.733	440	CC _{LOW}	−0.091	1.182	−0.06
	600	0.800	920	CC _{MID}	0.042	—	0.03
	650	0.867	1590	CC _{HIGH}	−0.077	0.846	−0.04
methanol (UA)	275	0.548	3.8 ₂	CC _{LOW}	−0.177	1.182	−0.09
	300	0.598	17 ₁	CC _{LOW}	0.081	0.625	0.06
				CC _{MID}	0.081	—	−0.01
	325	0.647	52 ₂	CC _{MID}	−0.050	—	−0.01
				CC _{HIGH}	−0.150	0.846	−0.09
	375	0.747	354 ₇	CC _{HIGH}	0.129	1.600	0.03
ethanol (UA)	275	0.535	1.4 ₂	CC _{LOW}	−0.208	1.182	−0.09
	300	0.584	8.1 ₅	CC _{MID}	0.095	—	0.03
	325	0.632	30 ₃	CC _{HIGH}	−0.176	0.846	−0.09
	375	0.730	234 ₁₀	CC _{LOW}	−0.216	1.267	−0.06
	425	0.827	1110 ₅₀	CC _{MID}	0.095	—	0.03
	475	0.924	3200 ₂₀₀	CC _{HIGH}	−0.170	0.789	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
propan-1-ol (UA)	300	0.558	4.5 ₂	CC _{LOW}	0.627	1.333	0.06
	350	0.651	43 ₁	CC _{LOW}	−0.653	1.286	−0.09
				CC _{MID}	−0.269	—	−0.01
				CC _{HIGH}	0.470	0.750	−0.05
	400	0.743	374 ₁₅	CC _{LOW}	0.114	1.250	0.05
				CC _{MID}	0.286	—	−0.02
				CC _{HIGH}	0.470	0.750	−0.05
	450	0.836	1210 ₃₀	CC _{LOW}	−0.545	3.815	−0.06
				CC _{MID}	−0.051	—	−0.02
				CC _{HIGH}	−0.508	0.778	−0.05
	500	0.929	3390 ₁₅₀	CC _{MID}	0.113	—	0.03
				CC _{HIGH}	0.091	0.800	−0.04
	515	0.957	3850 ₁₉₀	CC _{HIGH}	−0.143	0.262	−0.04
pentan-1-ol (UA)	300	0.518	0.46 ₄	CC _{LOW}	−0.204	1.333	−0.09
	350	0.604	10.3 ₄	CC _{LOW}	−0.253	1.286	−0.09
				CC _{MID}	0.087	—	−0.01
				CC _{HIGH}	−0.153	0.750	−0.09
	400	0.691	91 ₄	CC _{LOW}	−0.105	1.250	−0.09
				CC _{MID}	0.111	—	−0.01
				CC _{HIGH}	−0.153	0.750	−0.09
	450	0.777	407 ₂₅	CC _{MID}	0.047	—	0.03
				CC _{HIGH}	−0.197	0.778	−0.05
				CC _{LOW}	−0.185	3.767	−0.06
	500	0.864	1240 ₃₀	CC _{LOW}	−0.185	3.767	−0.06
				CC _{HIGH}	−0.084	0.800	−0.04
	550	0.950	2940 ₁₄₀	CC _{MID}	0.039	—	0.03
				CC _{HIGH}	−0.048	0.818	−0.04
	565	0.976	3520 ₁₀₀	CC _{HIGH}	−0.049	0.265	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
octan-1-ol (UA)	300	0.477	0.018 ₂	CC _{LOW}	−0.184	1.333	−0.09
	350	0.556	0.8 ₁	CC _{LOW}	−0.222	1.286	−0.09
				CC _{MID}	0.079	—	−0.01
	400	0.636	12 ₁	CC _{LOW}	−0.287	1.250	−0.09
				CC _{MID}	0.097	—	−0.01
				CC _{HIGH}	−0.138	0.750	−0.09
	450	0.715	83 ₃	CC _{LOW}	−0.144	1.222	−0.06
				CC _{MID}	0.127	—	−0.02
				CC _{HIGH}	−0.172	0.778	−0.05
	500	0.795	310 ₁₀	CC _{LOW}	−0.114	1.200	−0.06
				CC _{MID}	0.065	—	−0.02
				CC _{HIGH}	−0.229	0.800	−0.05
	550	0.874	810 ₄₀	CC _{MID}	0.052	—	0.03
				CC _{HIGH}	−0.118	0.818	−0.04
	600	0.954	1640 ₇₀	CC _{HIGH}	−0.095	0.833	−0.04
propan-2-ol (UA)	300	0.598	4.0 ₅	CC _{LOW}	−0.974	1.333	−0.09
	350	0.697	98 ₉	CC _{LOW}	−0.094	1.286	−0.09
				CC _{MID}	0.417	—	−0.01
	400	0.797	520 ₄₀	CC _{LOW}	−0.096	1.531	−0.06
				CC _{MID}	0.041	—	−0.02
				CC _{HIGH}	−0.730	0.750	−0.05
	450	0.896	1770 ₆₀	CC _{MID}	0.038	—	0.03
				CC _{HIGH}	−0.073	0.778	−0.04
	490	0.976	3700 ₅₀₀	CC _{HIGH}	−0.063	0.653	−0.04
butan-2-ol (UA)	300	0.573	3.4 ₁	CC _{LOW}	0.139	1.333	0.06
	350	0.668	43 ₂	CC _{LOW}	−0.408	1.286	−0.09
				CC _{MID}	−0.059	—	−0.01
	400	0.763	320 ₂₀	CC _{LOW}	0.080	1.250	0.05
				CC _{MID}	0.178	—	−0.02
				CC _{HIGH}	0.104	0.750	−0.05
	450	0.859	1110 ₅₀	CC _{MID}	−0.035	—	−0.03
				CC _{HIGH}	−0.317	0.778	−0.04
	500	0.954	3200 ₁₀₀	CC _{HIGH}	0.064	0.800	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
2-methylpropan-2-ol (UA)	300	0.595	6.3 ₃	CC _{LOW}	−0.532	1.333	−0.09
	350	0.694	113 ₈	CC _{LOW}	−0.537	1.286	−0.09
				CC _{MID}	0.228	—	−0.01
	400	0.794	661 ₂₃	CC _{LOW}	0.576	2.000	0.05
				CC _{MID}	0.235	—	−0.02
				CC _{HIGH}	−0.399	0.750	−0.05
	450	0.893	1720 ₁₅	CC _{MID}	−0.192	—	−0.03
				CC _{HIGH}	−0.418	0.778	−0.04
	480	0.952	3700 ₂₀₀	CC _{HIGH}	0.288	0.500	0.04
1,2-ethanediol (UA)	450	0.627	55 ₃	CC _{LOW}	−0.515	1.222	−0.09
	500	0.696	272 ₁₁	CC _{LOW}	0.382	1.200	0.06
				CC _{MID}	0.232	—	−0.01
	550	0.766	660 ₄₀	CC _{LOW}	−0.207	1.182	−0.06
				CC _{MID}	−0.174	—	−0.02
				CC _{HIGH}	−0.421	0.818	−0.05
	600	0.836	1900 ₃₀	CC _{MID}	0.095	—	0.03
				CC _{HIGH}	0.319	0.833	−0.05
	650	0.905	3900 ₂₀₀	CC _{HIGH}	−0.176	0.846	−0.04
1,3-propanediol (UA)	450	0.620	25.1 ₁₃	CC _{LOW}	−0.541	1.222	−0.09
	500	0.689	166 ₄	CC _{MID}	0.244	—	0.03
	550	0.758	500 ₃₀	CC _{MID}	−0.024	—	−0.02
				CC _{HIGH}	−0.443	0.818	−0.05
	600	0.826	1310 ₈₀	CC _{HIGH}	0.044	0.833	0.03
pentane-1,5-diol (UA)	500	0.704	77 ₅	CC _{LOW}	−0.184	1.200	−0.06
	550	0.775	320 ₂₀	CC _{MID}	0.083	—	0.03
	600	0.845	900 ₂₀	CC _{HIGH}	−0.153	0.833	−0.05
dimethyl ether (UA)	273	0.687	308 ₃	CC _{LOW}	−0.176	1.363	−0.09
	303	0.761	839 ₃₉	CC _{MID}	0.075	—	0.03
	329	0.827	1538 ₂₄	CC _{HIGH}	−0.129	0.734	−0.05

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
ethyl methyl ether (UA)	273	0.628	105 ₂	CC _{LOW}	−0.199	2.273	−0.09
	323	0.743	543 ₂₄	CC _{LOW}	0.161	1.987	0.05
				CC _{MID}	0.061	—	−0.02
	351	0.808	1025 ₃₀	CC _{MID}	−0.054	—	−0.02
				CC _{HIGH}	−0.087	0.440	−0.05
	367	0.845	1530 ₅₆	CC _{LOW}	−0.103	2.542	−0.06
				CC _{HIGH}	0.081	0.503	−0.05
	392	0.902	2472 ₄₆	CC _{HIGH}	−0.069	1.371	−0.04
	403	0.927	2867 ₅₀	CC _{HIGH}	−0.041	0.393	−0.04
diethyl ether (UA)	343	0.736	401 ₂₇	CC _{LOW}	0.082	2.055	0.05
	393	0.844	1202 ₂₃	CC _{MID}	−0.027	—	−0.02
dipropyl ether (UA)	325	0.617	40.6 ₂₆	CC _{LOW}	−0.219	1.308	−0.09
	375	0.712	209 ₄	CC _{LOW}	0.186	1.267	0.05
				CC _{MID}	0.095	—	−0.02
	425	0.806	619 ₁₈	CC _{MID}	−0.082	—	−0.02
				CC _{HIGH}	−0.167	0.765	−0.05
	475	0.901	1690 ₃₆	CC _{HIGH}	0.147	0.789	0.04
diisopropyl ether (UA)	350	0.706	184 ₈	CC _{LOW}	0.097	2.429	0.05
	400	0.806	622 ₁₉	CC _{MID}	−0.028	—	−0.02
	425	0.857	1069 ₅₂	CC _{HIGH}	0.040	0.412	0.04
methyl t-butyl ether (UA)	323	0.652	119 ₂	CC _{LOW}	−0.213	0.980	−0.09
				CC _{MID}	−0.015	—	−0.01
	373	0.752	511 ₁₁	CC _{MID}	0.108	—	0.03
				CC _{HIGH}	0.039	1.598	−0.05
	443	0.893	1820 ₃₆	CC _{HIGH}	−0.217	1.021	−0.04
1,2-dimethoxyethane (UA)	325	0.600	28.4 ₇	CC _{LOW}	−0.235	1.308	−0.09
	375	0.692	185 ₄	CC _{MID}	0.102	—	0.03
	425	0.784	648 ₂₅	CC _{MID}	0.038	—	0.03
				CC _{HIGH}	−0.179	0.765	−0.05
	475	0.876	1629 ₄₄	CC _{HIGH}	−0.068	0.789	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
2-methoxyethan-1-ol (UA)	347	0.589	18.5 ₆	CC _{LOW}	−0.474	1.237	−0.09
	397	0.674	150 ₃	CC _{MID}	0.212	—	0.03
	450	0.763	555 ₂₀	CC _{HIGH}	−0.383	0.808	−0.05
ethanal (UA)	260	0.561	18.5	CC _{LOW}	−0.288	1.308	−0.09
	300	0.647	135	CC _{MID}	0.125	—	0.03
	340	0.734	495	CC _{HIGH}	−0.221	0.765	−0.05
pentanal (UA)	300	0.530	6.07	CC _{LOW}	−0.106	1.267	−0.09
	340	0.601	35.7	CC _{MID}	0.047	—	0.03
	380	0.672	133	CC _{MID}	0.032	—	0.03
	420	0.742	364	CC _{HIGH}	−0.058	0.810	−0.05
2-pentanone (UA)	340	0.606	39.7	CC _{MID}	0.034	—	0.03
2-octanone (UA)	360	0.562	4.43	CC _{LOW}	−0.383	1.222	−0.09
	400	0.625	29.7	CC _{LOW}	−0.105	1.200	−0.09
				CC _{MID}	0.172	—	−0.01
	440	0.688	103	CC _{MID}	0.048	—	0.03
				CC _{HIGH}	−0.313	0.818	−0.09
	480	0.750	266	CC _{MID}	0.032	—	0.03
				CC _{HIGH}	−0.088	0.833	−0.05
	520	0.812	560	CC _{HIGH}	−0.058	0.846	−0.05
acetonitrile (UA)	298	0.543	13 ₁	CC _{LOW}	0.094	1.336	0.06
	348	0.634	81 ₁	CC _{MID}	−0.040	—	−0.01
	398	0.725	342 ₁₅	CC _{LOW}	−0.068	1.251	−0.06
				CC _{HIGH}	0.071	0.749	−0.05
	448	0.816	1060 ₃₀	CC _{MID}	0.030	—	0.03
	498	0.907	2480 ₆₀	CC _{HIGH}	−0.054	0.799	−0.04
propionitrile (UA)	298	0.536	8.7 ₂	CC _{LOW}	−0.183	0.752	−0.09
	348	0.626	69 ₂	CC _{MID}	0.105	—	0.03
	448	0.806	850 ₂₀	CC _{HIGH}	−0.244	1.330	−0.05

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
pyridine (UA)	298	0.482	1.2 ₁	CC _{LOW}	−0.466	1.336	−0.09
	348	0.563	22 ₂	CC _{LOW}	−0.206	0.961	−0.09
				CC _{MID}	0.200	—	−0.01
	398	0.644	137 ₅	CC _{LOW}	−0.121	1.672	−0.09
				CC _{MID}	0.105	—	−0.01
				CC _{HIGH}	−0.349	0.749	−0.09
	468	0.757	742 ₂₁	CC _{MID}	0.045	—	0.03
				CC _{HIGH}	−0.215	1.041	−0.05
	523	0.846	1895 ₂₁	CC _{HIGH}	−0.072	0.598	−0.05
pyrimidine (UA)	298	0.489	1.86 ₃	CC _{LOW}	0.288	2.712	0.06
	365	0.598	33 ₁	CC _{MID}	−0.078	—	−0.01
	398	0.652	106 ₉	CC _{LOW}	−0.155	1.266	−0.09
				CC _{HIGH}	0.106	0.369	−0.09
	440	0.721	356 ₂₁	CC _{MID}	0.068	—	0.03
	480	0.787	820 ₃₂	CC _{HIGH}	−0.122	0.790	−0.05
nitrobenzene (UA)	298	0.399	0.1 ₅	CC _{LOW}	0.060	1.147	0.06
	375	0.502	4 ₁	CC _{LOW}	0.142	2.422	0.06
				CC _{MID}	−0.028	—	−0.01
	484	0.648	105 ₆	CC _{MID}	−0.041	—	−0.01
				CC _{HIGH}	0.053	0.872	−0.09
	550	0.736	429 ₃₂	CC _{LOW}	−0.128	1.273	−0.06
				CC _{HIGH}	0.058	0.413	−0.05
	625	0.837	1409 ₄₆	CC _{MID}	0.056	—	0.03
methanethiol (UA)	700	0.937	3244 ₆₈	CC _{HIGH}	−0.100	0.786	−0.04
	220	0.463	4	CC _{LOW}	−0.271	1.182	−0.09
	240	0.505	17	CC _{MID}	0.124	—	0.03
ethanethiol (UA)	260	0.547	46	CC _{HIGH}	−0.229	0.846	−0.09
	240	0.478	5	CC _{LOW}	0.072	1.167	0.06
	260	0.518	14	CC _{MID}	−0.033	—	−0.01
	280	0.558	36	CC _{HIGH}	0.062	0.857	0.05

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
pentanethiol (UA)	300	0.498	10	CC _{LOW}	0.164	1.133	0.06
	320	0.532	16	CC _{LOW}	0.093	1.125	0.06
				CC _{MID}	−0.077	—	−0.01
	340	0.565	28	CC _{MID}	−0.044	—	−0.01
				CC _{HIGH}	0.145	0.882	−0.09
	360	0.598	50	CC _{MID}	−0.019	—	−0.01
				CC _{HIGH}	0.082	0.889	−0.09
octanethiol (UA)	380	0.566	8	CC _{MID}	0.037	—	0.03
2-butanethiol (UA)	280	0.503	4	CC _{LOW}	−0.149	1.143	−0.09
	300	0.539	13	CC _{MID}	0.070	—	0.03
	320	0.575	32	CC _{MID}	0.038	—	0.03
				CC _{HIGH}	−0.131	0.875	−0.09
2-methyl-1-propanethiol (UA)	300	0.539	6	CC _{LOW}	−0.335	1.133	−0.09
	320	0.575	23	CC _{LOW}	−0.120	1.125	−0.09
				CC _{MID}	0.157	—	−0.01
	340	0.610	56	CC _{MID}	0.057	—	0.03
				CC _{HIGH}	−0.296	0.882	−0.09
	360	0.646	111	CC _{HIGH}	−0.107	0.889	−0.09
dimethyl disulfide (UA)	300	0.495	4	CC _{LOW}	−0.269	1.133	−0.09
	320	0.528	13	CC _{MID}	0.126	—	0.03
	340	0.561	29	CC _{HIGH}	−0.238	0.882	−0.09
thiophene (UA)	293	0.484	8	CC _{LOW}	−0.474	1.341	−0.09
	343	0.567	74	CC _{MID}	0.202	—	0.03
	393	0.650	273	CC _{HIGH}	−0.353	0.746	−0.09
	443	0.732	755	CC _{LOW}	0.199	1.617	0.05
	493	0.815	1632	CC _{LOW}	−0.210	1.666	−0.06
				CC _{MID}	−0.076	—	−0.02
	530	0.876	2973	CC _{MID}	0.079	—	0.03
				CC _{HIGH}	0.123	0.619	−0.04
	555	0.917	3757	CC _{HIGH}	−0.126	0.600	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
1,3-butadiene (UA)	275	0.644	170 ₂	CC _{MID}	−0.017	—	−0.01
	300	0.703	386 ₄	CC _{MID}	0.037	—	0.03
				CC _{HIGH}	0.032	0.833	−0.05
	325	0.761	722 ₇	CC _{HIGH}	−0.068	0.846	−0.05
methyl acrylate (UA)	298	0.530	11.2 ₅	CC _{LOW}	0.168	4.933	0.06
	340	0.605	61 ₁	CC _{MID}	−0.028	—	−0.01
	350	0.622	89 ₁	CC _{MID}	0.038	—	0.03
	360	0.640	118 ₂	CC _{MID}	−0.016	—	−0.01
	400	0.711	341 ₃	CC _{HIGH}	0.074	3.500	0.03
ethyl acrylate (UA)	298	0.516	4.9 ₂	CC _{LOW}	−0.253	7.698	−0.09
	380	0.657	115 ₄	CC _{LOW}	−0.224	1.842	−0.09
				CC _{MID}	−0.029	—	−0.01
	450	0.779	652 ₁₅	CC _{LOW}	0.119	2.333	0.05
				CC _{MID}	0.079	—	−0.02
				CC _{HIGH}	0.195	5.756	−0.05
	500	0.865	1481 ₃₂	CC _{MID}	−0.036	—	−0.03
				CC _{HIGH}	−0.121	0.543	−0.04
n-butyl acrylate (UA)	525	0.908	2215 ₅₃	CC _{HIGH}	0.051	0.429	0.04
	450	0.721	196 ₂	CC _{LOW}	−0.159	1.222	−0.06
				CC _{MID}	−0.026	—	−0.02
	500	0.801	570 ₁₂	CC _{LOW}	0.183	2.300	0.05
				CC _{MID}	0.071	—	−0.02
				CC _{HIGH}	0.070	1.700	−0.05
	550	0.881	1199 ₂₉	CC _{MID}	−0.056	—	−0.03
				CC _{HIGH}	−0.130	0.818	−0.04
n-octyl acrylate (UA)	575	0.921	1794 ₁₄	CC _{HIGH}	0.080	0.435	0.04
	450	0.640	23.7 ₈	CC _{LOW}	0.092	4.444	0.06
	490	0.697	68 ₁	CC _{MID}	−0.017	—	−0.01
	550	0.783	261 ₄	CC _{HIGH}	−0.113	3.636	−0.05

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
2-ethylhexyl acrylate (UA)	450	0.650	32.6 ₉	CC _{LOW}	−0.130	3.267	−0.09
	480	0.694	75 ₂	CC _{MID}	0.031	—	0.03
	490	0.708	93 ₂	CC _{MID}	−0.020	—	−0.02
	500	0.723	119 ₅	CC _{LOW}	−0.113	1.200	−0.06
				CC _{HIGH}	0.040	0.960	−0.05
	550	0.795	328 ₆	CC _{LOW}	0.150	1.182	0.05
				CC _{MID}	0.051	—	−0.02
				CC _{HIGH}	−0.084	4.455	−0.05
	600	0.867	695 ₃₂	CC _{MID}	−0.069	—	−0.03
				CC _{HIGH}	−0.094	0.833	−0.04
	650	0.939	1490 ₂₁	CC _{HIGH}	0.127	0.846	0.04
2-hydroxyethyl acrylate (UA)	470	0.693	128 ₂	CC _{MID}	−0.014	—	−0.01
	500	0.737	273 ₅	CC _{HIGH}	−0.060	1.880	−0.05
	600	0.885	1709 ₂₅	CC _{HIGH}	−0.040	0.833	−0.04
methyl methacrylate (UA)	298	0.508	4.7 ₂	CC _{LOW}	0.210	7.698	0.06
	360	0.614	57 ₁	CC _{MID}	−0.024	—	−0.01
ethyl methacrylate (UA)	360	0.602	35 ₁	CC _{LOW}	0.131	2.167	0.06
	380	0.636	63.1 ₉	CC _{MID}	−0.041	—	−0.01
	390	0.652	88 ₄	CC _{HIGH}	0.061	0.462	0.05
n-butyl methacrylate (UA)	375	0.586	14.2 ₅	CC _{LOW}	0.068	1.400	0.06
	400	0.625	32 ₁	CC _{MID}	−0.028	—	−0.01
	420	0.657	60 ₂	CC _{MID}	−0.014	—	−0.01
	440	0.688	111 ₂	CC _{MID}	0.049	—	0.03
	500	0.782	392 ₈	CC _{HIGH}	−0.301	5.160	−0.05
	600	0.938	1818 ₃₁	CC _{HIGH}	−0.044	0.833	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
cyclopentane (UA)	272	0.528	23.0 ₇	CC _{LOW}	0.170	1.213	0.06
	301	0.584	66 ₂	CC _{LOW}	-0.172	1.193	-0.09
				CC _{MID}	-0.077	—	-0.01
	330	0.641	181 ₄	CC _{LOW}	0.136	1.590	0.06
				CC _{MID}	0.079	—	-0.01
				CC _{HIGH}	0.140	0.824	-0.09
	359	0.697	365 ₇	CC _{LOW}	-0.092	1.069	-0.09
				CC _{MID}	-0.052	—	-0.01
				CC _{HIGH}	-0.144	0.838	-0.09
	380	0.738	618 ₄	CC _{LOW}	0.053	1.609	0.05
				CC _{MID}	0.044	—	-0.02
				CC _{HIGH}	0.086	0.629	-0.05
	402	0.781	928 ₁₃	CC _{MID}	-0.020	—	-0.02
				CC _{HIGH}	-0.086	0.936	-0.05
cyclohexane (UA)	417	0.810	1235 ₂₀	CC _{HIGH}	0.033	0.621	0.03
	482	0.936	3260 ₃₀	CC _{HIGH}	0.041	1.007	0.04
	323	0.583	52.9 ₁₀	CC _{MID}	-0.026	—	-0.01
tetrahydrofuran (UA)	330	0.607	99 ₂	CC _{LOW}	0.065	1.121	0.06
				CC _{MID}	0.041	—	-0.01
	350	0.643	171 ₂	CC _{MID}	-0.031	—	-0.01
	370	0.680	295 ₂	CC _{MID}	-0.012	—	-0.01
				CC _{HIGH}	0.058	0.892	-0.09
1,3-dioxolane (UA)	290	0.513	11.0 ₂	CC _{LOW}	0.135	1.138	0.06
	310	0.549	25.7 ₁₄	CC _{MID}	-0.063	—	-0.01
	330	0.584	61 ₂	CC _{LOW}	-0.107	1.121	-0.09
				CC _{MID}	0.034	—	-0.01
				CC _{HIGH}	0.119	0.879	-0.09
	350	0.619	123 ₂	CC _{LOW}	0.070	1.114	0.06
				CC _{MID}	0.050	—	-0.01
	370	0.655	209 ₃	CC _{MID}	-0.033	—	-0.01
				CC _{HIGH}	-0.095	0.892	-0.09
	390	0.690	358 ₄	CC _{HIGH}	0.062	0.897	0.05

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
oxane (UA)	340	0.582	54.1 ₁₃	CC _{MID}	0.034	—	0.03
	360	0.616	98 ₂	CC _{MID}	−0.016	—	−0.01
	400	0.685	276 ₃	CC _{MID}	−0.015	—	−0.01
	440	0.753	634 ₄	CC _{MID}	−0.022	—	−0.02
	460	0.788	930 ₇	CC _{HIGH}	0.041	0.913	0.03
1,3-dioxane (UA)	335	0.552	25.5 ₆	CC _{MID}	−0.012	—	−0.01
	355	0.585	52.5 ₁₂	CC _{MID}	−0.015	—	−0.01
	375	0.618	103 ₂	CC _{MID}	0.040	—	0.03
1,4-dioxane (UA)	355	0.604	85 ₂	CC _{MID}	−0.010	—	−0.01
	555	0.944	4040 ₆₀	CC _{HIGH}	0.049	0.928	0.04
1,3,5-trioxane (UA)	510	0.839	1930 ₂₀	CC _{LOW}	0.051	1.078	0.05
	550	0.905	3460 ₃₀	CC _{HIGH}	0.047	0.927	0.04
	570	0.938	4370 ₃₀	CC _{HIGH}	−0.050	0.930	−0.04
SPC-pol3 (pol)	373	0.602	155 ₆	CC _{LOW}	0.195	1.268	0.06
	423	0.682	638 ₂₇	CC _{LOW}	−0.114	1.236	−0.09
				CC _{MID}	−0.086	—	−0.01
	473	0.763	2270 ₉₀	CC _{MID}	0.051	—	0.03
				CC _{HIGH}	0.153	0.789	−0.05
	523	0.844	5780 ₁₆₀	CC _{HIGH}	−0.092	0.809	−0.05
TIP4P-pol1 (pol)	373	0.643	173 ₁₃	CC _{LOW}	0.073	1.268	0.06
	423	0.729	734 ₃₇	CC _{LOW}	0.064	1.236	0.05
				CC _{MID}	−0.032	—	−0.02
	473	0.816	2430 ₇₀	CC _{MID}	−0.029	—	−0.02
				CC _{HIGH}	0.057	0.789	−0.05
	523	0.902	6740 ₃₇₀	CC _{HIGH}	0.052	0.809	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
TIP4P-pol3 (pol)	373	0.643	275 ₁₃	CC _{LOW}	0.086	1.268	0.06
	423	0.729	1122 ₄₈	CC _{LOW}	−0.092	1.236	−0.06
				CC _{MID}	−0.038	—	−0.02
	473	0.816	3640 ₁₀₀	CC _{MID}	0.041	—	0.03
				CC _{HIGH}	0.068	0.789	−0.05
	523	0.902	8750 ₂₆₀	CC _{HIGH}	−0.075	0.809	−0.04
methane (EH)	120	0.633	206 ₇	CC _{LOW}	0.162	1.250	0.06
	135	0.712	501 ₁₅	CC _{LOW}	−0.152	1.222	−0.06
				CC _{MID}	−0.072	—	−0.02
	150	0.791	1161 ₃₈	CC _{MID}	0.068	—	0.03
				CC _{HIGH}	0.129	0.800	−0.05
	165	0.870	2040 ₆₀	CC _{HIGH}	−0.124	0.818	−0.04
ethane (EH)	185	0.601	101 ₃	CC _{LOW}	−0.415	2.486	−0.09
	215	0.698	447 ₁₃	CC _{LOW}	0.144	1.140	0.06
				CC _{MID}	0.119	—	−0.01
	230	0.747	688 ₃₁	CC _{MID}	−0.067	—	−0.02
				CC _{HIGH}	−0.167	0.402	−0.05
	245	0.795	1140 ₆₀	CC _{HIGH}	0.127	0.878	0.03
propane (EH)	260	0.699	302 ₁₁	CC _{MID}	−0.018	—	−0.01
	280	0.753	582 ₂₃	CC _{LOW}	0.083	1.143	0.05
				CC _{HIGH}	0.033	0.857	−0.05
	300	0.806	973 ₄₃	CC _{MID}	−0.039	—	−0.02
				CC _{HIGH}	−0.055	0.867	−0.05
	320	0.860	1641 ₅₅	CC _{HIGH}	0.073	0.875	0.04
n-pentane (EH)	360	0.761	461 ₃₆	CC _{LOW}	0.065	1.167	0.05
	390	0.825	864 ₄₆	CC _{MID}	−0.030	—	−0.02
	420	0.888	1565 ₈₇	CC _{LOW}	0.231	2.143	0.06
				CC _{HIGH}	0.056	0.857	−0.04
	440	0.930	2230 ₁₃₀	CC _{MID}	−0.073	—	−0.03
	450	0.951	2930 ₈₀	CC _{HIGH}	0.108	0.467	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
n-octane (EH)	400	0.697	105 ₈	CC _{LOW}	−0.236	1.567	−0.09
	440	0.767	306 ₁₂	CC _{LOW}	0.100	1.136	0.05
				CC _{MID}	0.092	—	−0.02
	470	0.819	521 ₂₂	CC _{MID}	−0.047	—	−0.02
				CC _{HIGH}	−0.151	0.638	−0.05
	500	0.871	909 ₃₈	CC _{HIGH}	0.088	0.880	0.04
	530	0.923	1424 ₃₂	CC _{HIGH}	−0.045	0.887	−0.04
n-dodecane (EH)	450	0.675	35 ₂	CC _{LOW}	0.101	1.222	0.06
	500	0.750	111 ₆	CC _{LOW}	0.156	1.671	0.05
				CC _{MID}	−0.045	—	−0.02
	550	0.825	310 ₁₄	CC _{MID}	−0.058	—	−0.02
				CC _{HIGH}	0.083	0.818	−0.05
	585	0.877	629 ₃₀	CC _{HIGH}	0.093	0.598	0.04
methylamine (EH)	267	0.622	99 ₂	CC _{LOW}	−0.139	0.781	−0.09
				CC _{MID}	−0.036	—	−0.01
	298	0.695	362 ₁₅	CC _{MID}	0.078	—	0.03
				CC _{HIGH}	0.090	1.530	−0.09
	350	0.816	1592 ₂₄	CC _{HIGH}	−0.178	1.280	−0.05
dimethylamine (EH)	250	0.564	20 ₂	CC _{LOW}	−0.165	1.320	−0.09
	280	0.632	104 ₆	CC _{LOW}	−0.093	0.833	−0.09
				CC _{MID}	0.071	—	−0.01
	308	0.695	320 ₁₇	CC _{MID}	0.051	—	0.03
				CC _{HIGH}	−0.125	0.758	−0.09
	350	0.790	1102 ₇₈	CC _{MID}	0.040	—	0.03
				CC _{HIGH}	−0.112	1.200	−0.05
	400	0.903	3173 ₉₉	CC _{HIGH}	−0.076	0.917	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
trimethylamine (EH)	240	0.554	25 ₁	CC _{LOW}	0.876	2.374	0.06
	275	0.635	98 ₂	CC _{LOW}	−0.397	0.402	−0.09
				CC _{MID}	−0.260	—	−0.01
	293	0.677	252 ₂₈	CC _{MID}	0.283	—	0.03
				CC _{HIGH}	0.369	0.421	−0.09
	350	0.808	985 ₅₅	CC _{MID}	−0.022	—	−0.02
				CC _{HIGH}	−0.987	2.488	−0.05
ethylamine (EH)	270	0.597	46 ₂	CC _{LOW}	0.479	0.523	0.06
	293	0.648	92 ₂	CC _{LOW}	−0.639	1.556	−0.09
				CC _{MID}	−0.315	—	−0.01
	350	0.774	865 ₄₆	CC _{MID}	0.250	—	0.03
				CC _{HIGH}	0.916	1.912	−0.05
	400	0.885	2422 ₁₄₅	CC _{HIGH}	−0.410	0.643	−0.04
diethylamine (EH)	293	0.584	40 ₃	CC _{LOW}	0.435	1.002	0.06
	329	0.655	98 ₆	CC _{LOW}	0.100	1.188	0.06
				CC _{MID}	−0.217	—	−0.01
	375	0.747	370 ₂₃	CC _{LOW}	−0.193	1.267	−0.06
				CC _{MID}	−0.046	—	−0.02
				CC _{HIGH}	0.434	0.998	−0.05
	425	0.847	1231 ₆₃	CC _{MID}	0.085	—	0.03
				CC _{HIGH}	0.084	0.841	−0.05
	475	0.946	2731 ₁₀₄	CC _{HIGH}	−0.152	0.789	−0.04
triethylamine (EH)	293	0.544	13.9 ₁	CC _{LOW}	0.316	1.638	0.06
	363	0.673	132 ₁	CC _{LOW}	−0.209	1.139	−0.09
				CC _{MID}	−0.120	—	−0.01
	425	0.788	633 ₅	CC _{MID}	0.098	—	0.03
				CC _{HIGH}	0.193	0.611	−0.05
	500	0.928	2088 ₂₃	CC _{HIGH}	−0.183	0.878	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
nitromethane (EH)	293	0.497	2.2 ₂	CC _{LOW}	−0.422	1.679	−0.09
	375	0.637	99 ₃	CC _{LOW}	−0.316	2.000	−0.09
				CC _{MID}	0.158	—	−0.01
	450	0.764	743 ₁₇	CC _{LOW}	0.235	1.222	0.05
				CC _{MID}	0.105	—	−0.02
				CC _{HIGH}	−0.251	0.596	−0.05
	500	0.849	1738 ₄₅	CC _{LOW}	−0.227	2.300	−0.06
				CC _{MID}	−0.106	—	−0.02
				CC _{HIGH}	−0.158	0.500	−0.05
	550	0.934	4221 ₄₈	CC _{MID}	0.069	—	0.03
				CC _{HIGH}	0.192	0.818	−0.04
	575	0.976	5624 ₁₉₀	CC _{HIGH}	−0.099	0.435	−0.04
nitroethane (EH)	323	0.555	9 ₁	CC _{LOW}	−0.094	1.461	−0.09
	388	0.667	108 ₆	CC _{MID}	0.038	—	0.03
	450	0.773	555 ₁₅	CC _{LOW}	0.056	1.222	0.05
				CC _{HIGH}	−0.064	0.685	−0.05
	550	0.945	3568 ₁₆₀	CC _{HIGH}	0.046	0.818	0.04
acetonitrile (EH)	348	0.636	81 ₁	CC _{LOW}	−0.153	1.287	−0.09
	398	0.728	367 ₆	CC _{LOW}	0.157	1.251	0.05
				CC _{MID}	0.067	—	−0.02
	448	0.819	1054 ₂₀	CC _{MID}	−0.070	—	−0.02
				CC _{HIGH}	−0.119	0.777	−0.05
	498	0.910	2776 ₆₂	CC _{HIGH}	0.125	0.799	0.04
propionitrile (EH)	348	0.627	56 ₁	CC _{LOW}	−0.125	1.287	−0.09
	398	0.717	266 ₅	CC _{MID}	0.054	—	0.03
	448	0.807	810 ₂₀	CC _{HIGH}	−0.097	0.777	−0.05
acetamide (EH)	400	0.535	3.4 ₁	CC _{LOW}	−0.218	1.250	−0.09
	450	0.602	27 ₁	CC _{MID}	0.097	—	0.03
	500	0.669	119 ₃	CC _{MID}	−0.016	—	−0.01
				CC _{HIGH}	−0.174	0.800	−0.09

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
propanamide (EH)	400	0.542	3.4 ₂	CC _{LOW}	−0.109	1.250	−0.09
	450	0.610	25 ₁	CC _{MID}	0.049	—	0.03
butanamide (EH)	400	0.535	1.6 ₁	CC _{LOW}	−0.115	1.250	−0.09
	450	0.602	15 ₁	CC _{MID}	0.051	—	0.03
	500	0.668	82 ₃	CC _{MID}	0.035	—	0.03
				CC _{HIGH}	−0.092	0.800	−0.09
	550	0.735	309 ₉	CC _{HIGH}	−0.063	0.818	−0.05
benzene (EH)	415	0.737	470 ₆₀	CC _{LOW}	0.060	1.139	0.05
	465	0.826	1250 ₁₁₀	CC _{MID}	−0.028	—	−0.02
	520	0.924	3110 ₁₅₀	CC _{HIGH}	0.053	0.878	0.04
pyridine (EH)	298	0.482	2 ₁	CC _{LOW}	−0.317	1.335	−0.09
	348	0.562	26 ₁	CC _{LOW}	−0.180	0.961	−0.09
				CC _{MID}	0.136	—	−0.01
	398	0.643	140 ₂₀	CC _{LOW}	0.078	1.672	0.06
				CC _{MID}	0.092	—	−0.01
				CC _{HIGH}	−0.237	0.749	−0.09
	468	0.756	670 ₅₀	CC _{LOW}	−0.111	1.299	−0.06
				CC _{MID}	−0.029	—	−0.02
				CC _{HIGH}	−0.187	1.041	−0.05
	523	0.845	1790 ₁₃₀	CC _{MID}	0.048	—	0.03
				CC _{HIGH}	0.047	0.598	−0.05
	575	0.929	3500 ₂₀₀	CC _{HIGH}	−0.086	0.770	−0.04
pyrimidine (EH)	298	0.472	0.3 ₄	CC _{LOW}	−1.528	1.335	−0.09
	348	0.551	18 ₅	CC _{LOW}	−0.452	0.961	−0.09
				CC _{MID}	0.654	—	−0.01
	398	0.630	123 ₁₆	CC _{LOW}	0.233	1.672	0.06
				CC _{MID}	0.231	—	−0.01
				CC _{HIGH}	−1.144	0.749	−0.09
	468	0.741	568 ₁₂	CC _{MID}	−0.087	—	−0.02
				CC _{HIGH}	−0.471	1.041	−0.05
	523	0.828	1630 ₈₀	CC _{HIGH}	0.139	0.598	0.03

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
pyrazine (EH)	348	0.565	20 ₄	CC _{LOW}	−0.407	0.961	−0.09
	398	0.646	150 ₂₄	CC _{MID}	0.208	—	0.03
	468	0.760	800 ₅₀	CC _{LOW}	0.054	1.581	0.05
				CC _{HIGH}	−0.424	1.041	−0.05
	523	0.849	2100 ₃₀₀	CC _{MID}	−0.021	—	−0.02
pyridazine (EH)	298	0.392	0.10 ₄	CC _{LOW}	2.382	1.335	0.06
	348	0.457	0.20 ₇	CC _{LOW}	0.551	0.961	0.06
				CC _{MID}	−1.020	—	−0.01
	398	0.523	2.0 ₃	CC _{LOW}	−0.489	1.672	−0.09
				CC _{MID}	−0.281	—	−0.01
				CC _{HIGH}	1.784	0.749	−0.09
	468	0.615	39 ₅	CC _{LOW}	−0.177	0.751	−0.09
				CC _{MID}	0.183	—	−0.01
				CC _{HIGH}	0.573	1.041	−0.09
	523	0.687	172 ₁₁	CC _{LOW}	0.073	2.490	0.06
				CC _{MID}	0.101	—	−0.01
				CC _{HIGH}	−0.292	0.598	−0.09
furan (EH)	620	0.815	980 ₉₀	CC _{MID}	−0.021	—	−0.02
				CC _{HIGH}	−0.236	1.331	−0.05
	250	0.506	8 ₁	CC _{LOW}	−0.413	1.400	−0.09
	300	0.607	98 ₇	CC _{LOW}	0.122	1.333	0.06
				CC _{MID}	0.172	—	−0.01
	350	0.709	437 ₁₁	CC _{LOW}	−0.092	1.571	−0.06
				CC _{MID}	−0.052	—	−0.02
				CC _{HIGH}	−0.295	0.714	−0.05
	400	0.810	1470 ₇₀	CC _{MID}	0.036	—	0.03
				CC _{HIGH}	0.092	0.750	−0.05
	440	0.891	3000 ₂₀₀	CC _{HIGH}	−0.059	0.636	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
thiophene (EH)	393	0.670	274 ₁₉	CC _{LOW}	−0.189	1.289	−0.09
				CC _{MID}	0.036	—	−0.01
	460	0.784	1200 ₁₀₀	CC _{MID}	0.083	—	0.03
				CC _{HIGH}	−0.073	1.043	−0.05
	530	0.903	3260 ₁₆₀	CC _{HIGH}	−0.147	0.776	−0.04
pyrrole (EH)	300	0.461	1.0 ₂	CC _{LOW}	−0.783	1.467	−0.09
	370	0.568	39 ₁₀	CC _{LOW}	0.109	1.230	0.06
				CC _{MID}	0.317	—	−0.01
	440	0.676	278 ₂₂	CC _{LOW}	−0.232	1.758	−0.09
				CC _{MID}	−0.049	—	−0.01
				CC _{HIGH}	−0.534	0.682	−0.09
	520	0.799	1500 ₁₀₀	CC _{MID}	0.084	—	0.03
				CC _{HIGH}	0.088	0.813	−0.05
thiazole (EH)	580	0.891	3430 ₈₀	CC _{HIGH}	−0.132	0.569	−0.04
	300	0.472	9.0 ₄	CC _{LOW}	1.418	1.333	0.06
	350	0.550	28 ₄	CC _{LOW}	−0.525	1.286	−0.09
				CC _{MID}	−0.608	—	−0.01
	400	0.629	190 ₂₇	CC _{MID}	0.230	—	0.03
				CC _{HIGH}	1.064	0.750	−0.09
	450	0.708	560 ₅₀	CC _{LOW}	0.075	1.222	0.05
				CC _{HIGH}	−0.408	0.778	−0.05
oxazole (EH)	500	0.786	1330 ₆₀	CC _{MID}	−0.034	—	−0.02
	550	0.865	2870 ₁₇₀	CC _{HIGH}	0.061	0.818	0.04
	300	0.557	43 ₄	CC _{LOW}	0.319	1.333	0.06
	350	0.649	195 ₁₃	CC _{LOW}	−0.114	1.286	−0.09
				CC _{MID}	−0.137	—	−0.01
	400	0.742	770 ₃₀	CC _{MID}	0.050	—	0.03
				CC _{HIGH}	0.240	0.750	−0.05
	450	0.835	2050 ₁₅₀	CC _{HIGH}	−0.089	0.778	−0.05

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
isoxazole (EH)	400	0.677	272 ₁₇	CC _{MID}	0.038	—	0.03
	450	0.761	835 ₂₄	CC _{HIGH}	−0.067	0.778	−0.05
	475	0.804	1340 ₉₀	CC _{MID}	−0.030	—	−0.02
	540	0.914	4100 ₃₀₀	CC _{HIGH}	0.093	2.167	0.04
imidazole (EH)	400	0.489	3.0 ₃	CC _{LOW}	0.121	1.250	0.06
	450	0.550	17 ₇	CC _{MID}	−0.054	—	−0.01
	500	0.611	75 ₄	CC _{MID}	0.035	—	0.03
				CC _{HIGH}	0.097	0.800	−0.09
pyrazole (EH)	350	0.475	3.0 ₂	CC _{LOW}	−0.522	1.286	−0.09
	400	0.543	28 ₆	CC _{LOW}	0.265	1.250	0.06
				CC _{MID}	0.228	—	−0.01
	450	0.611	106 ₆	CC _{LOW}	−0.157	1.222	−0.09
				CC _{MID}	−0.118	—	−0.01
				CC _{HIGH}	−0.406	0.778	−0.09
	500	0.678	380 ₃₀	CC _{MID}	0.071	—	0.03
				CC _{HIGH}	0.212	0.800	−0.09
fluorobenzene (EH)	550	0.746	950 ₆₀	CC _{HIGH}	−0.128	0.818	−0.05
	300	0.531	9.6 ₈	CC _{LOW}	−0.317	1.331	−0.09
	350	0.619	80 ₄	CC _{MID}	0.136	—	0.03
	400	0.708	310 ₃₀	CC _{HIGH}	−0.238	0.751	−0.05
chlorobenzene (EH)	298	0.471	5.7 ₂	CC _{LOW}	1.138	1.335	0.06
	348	0.550	17.6 ₁₀	CC _{LOW}	−0.213	0.961	−0.09
				CC _{MID}	−0.487	—	−0.01
	398	0.629	96 ₈	CC _{LOW}	0.229	1.672	0.06
				CC _{MID}	0.108	—	−0.01
				CC _{HIGH}	0.852	0.749	−0.09
	468	0.739	450 ₃₀	CC _{LOW}	−0.064	1.299	−0.06
				CC _{MID}	−0.086	—	−0.02
				CC _{HIGH}	−0.221	1.041	−0.05
	523	0.826	1300 ₇₀	CC _{HIGH}	0.137	0.598	0.03
	575	0.909	2800 ₂₀₀	CC _{HIGH}	−0.049	0.770	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
bromobenzene (EH)	300	0.447	3.6 ₁	CC _{LOW}	2.151	1.400	0.06
	360	0.537	8.9 ₂	CC _{LOW}	−0.316	1.335	−0.09
				CC _{MID}	−0.896	—	−0.01
	420	0.626	79 ₁₄	CC _{MID}	0.135	—	0.03
				CC _{HIGH}	1.537	0.714	−0.09
	480	0.715	320 ₃₀	CC _{HIGH}	−0.236	0.749	−0.05
1,2-dichlorobenzene (EH)	348	0.498	9 ₂	CC _{LOW}	1.236	2.050	0.06
	398	0.570	19 ₂	CC _{MID}	−0.405	—	−0.01
	428	0.613	50 ₂	CC _{LOW}	−0.244	0.889	−0.09
				CC _{MID}	−0.031	—	−0.01
				CC _{HIGH}	0.603	0.488	−0.09
	468	0.670	160 ₂₀	CC _{LOW}	0.174	1.299	0.06
				CC _{MID}	0.129	—	−0.01
				CC _{HIGH}	0.066	1.134	−0.09
	523	0.748	450 ₄₀	CC _{LOW}	−0.068	0.862	−0.06
				CC _{MID}	−0.076	—	−0.02
				CC _{HIGH}	−0.275	1.125	−0.05
	575	0.823	1140 ₅₀	CC _{MID}	0.036	—	0.03
				CC _{HIGH}	0.134	0.770	−0.05
	650	0.930	3100 ₂₀₀	CC _{HIGH}	−0.078	1.161	−0.04
1,3-dichlorobenzene (EH)	398	0.575	27 ₃	CC _{LOW}	0.404	0.882	0.06
	428	0.618	53 ₅	CC _{LOW}	−0.230	0.889	−0.09
				CC _{MID}	−0.215	—	−0.01
	468	0.676	180 ₂₀	CC _{LOW}	−0.125	1.299	−0.09
				CC _{MID}	0.122	—	−0.01
				CC _{HIGH}	0.458	1.134	−0.09
	523	0.755	550 ₃₀	CC _{LOW}	0.069	0.862	0.05
				CC _{MID}	0.054	—	−0.02
				CC _{HIGH}	−0.259	1.125	−0.05
	575	0.830	1180 ₈₀	CC _{MID}	−0.037	—	−0.02
				CC _{HIGH}	−0.096	0.770	−0.05
	650	0.938	3100 ₁₀₀	CC _{HIGH}	0.080	1.161	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
1,4-dichlorobenzene (EH)	348	0.503	3.9 ₇	CC _{LOW}	−0.174	2.050	−0.09
	398	0.575	22 ₄	CC _{LOW}	0.228	0.882	0.06
				CC _{MID}	0.057	—	−0.01
	428	0.619	47 ₆	CC _{MID}	−0.121	—	−0.01
	468	0.677	144 ₅	CC _{MID}	0.036	—	0.03
				CC _{HIGH}	0.259	1.134	−0.09
	523	0.756	470 ₃₀	CC _{LOW}	−0.132	0.862	−0.06
				CC _{HIGH}	−0.077	1.125	−0.05
	575	0.831	1190 ₉₀	CC _{MID}	0.071	—	0.03
	650	0.940	3000 ₈₀	CC _{HIGH}	−0.153	1.161	−0.04
1,2,3-trichlorobenzene (EH)	400	0.532	6 ₄	CC _{LOW}	−0.535	2.375	−0.09
	450	0.598	38 ₆	CC _{MID}	0.158	—	0.03
	475	0.632	66 ₉	CC _{LOW}	0.185	1.105	0.06
				CC _{MID}	−0.012	—	−0.01
				CC _{HIGH}	−0.225	0.421	−0.09
	500	0.665	111 ₁₁	CC _{MID}	−0.088	—	−0.01
	525	0.698	210 ₁₄	CC _{LOW}	−0.168	0.571	−0.09
				CC _{HIGH}	0.167	0.905	−0.09
	550	0.731	380 ₃₀	CC _{LOW}	0.214	1.182	0.05
				CC _{MID}	0.107	—	−0.02
	600	0.798	800 ₈₀	CC _{LOW}	−0.295	1.167	−0.06
				CC _{MID}	−0.098	—	−0.02
				CC _{HIGH}	−0.293	1.750	−0.05
	650	0.864	1800 ₂₀₀	CC _{LOW}	0.239	2.231	0.06
				CC _{MID}	0.136	—	−0.03
				CC _{HIGH}	0.181	0.846	−0.04
	700	0.931	2800 ₂₀₀	CC _{MID}	−0.074	—	−0.03
				CC _{HIGH}	−0.253	0.857	−0.04
	725	0.964	3800 ₁₀₀	CC _{HIGH}	0.107	0.448	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
1,3,5-trichlorobenzene (EH)	400	0.546	9 ₂	CC _{LOW}	0.505	2.375	0.06
	450	0.614	37 ₇	CC _{LOW}	−0.227	1.111	−0.09
				CC _{MID}	−0.150	—	−0.01
	475	0.648	83 ₁₃	CC _{LOW}	0.073	1.105	0.06
				CC _{MID}	0.108	—	−0.01
				CC _{HIGH}	0.213	0.421	−0.09
	500	0.682	140 ₂₀	CC _{MID}	−0.035	—	−0.01
				CC _{HIGH}	−0.204	0.900	−0.09
	525	0.716	240 ₃₀	CC _{LOW}	0.052	0.571	0.05
				CC _{HIGH}	0.066	0.905	−0.05
	550	0.750	380 ₃₀	CC _{LOW}	−0.115	1.182	−0.06
				CC _{MID}	−0.033	—	−0.02
	600	0.819	930 ₈₀	CC _{MID}	0.053	—	0.03
				CC _{HIGH}	0.091	1.750	−0.05
	650	0.887	1800 ₂₀₀	CC _{HIGH}	−0.097	0.846	−0.04
hexachlorobenzene (EH)	575	0.648	60 ₂₀	CC _{LOW}	−0.273	1.087	−0.09
				CC _{MID}	−0.019	—	−0.01
	600	0.676	114 ₁₁	CC _{MID}	0.131	—	0.03
	625	0.705	160 ₃₀	CC _{LOW}	0.094	0.560	0.05
				CC _{MID}	−0.026	—	−0.02
				CC _{HIGH}	−0.252	0.920	−0.05
	650	0.733	230 ₂₀	CC _{LOW}	−0.108	1.154	−0.06
				CC _{MID}	−0.060	—	−0.02
				CC _{HIGH}	0.050	0.923	−0.05
	700	0.789	520 ₆₀	CC _{LOW}	0.080	1.143	0.05
				CC _{MID}	0.050	—	−0.02
				CC _{HIGH}	0.168	1.786	−0.05
	750	0.846	960 ₇₀	CC _{LOW}	−0.115	2.200	−0.06
				CC _{MID}	−0.037	—	−0.02
				CC _{HIGH}	−0.094	0.867	−0.05
	800	0.902	1760 ₁₅₀	CC _{MID}	0.036	—	0.03
				CC _{HIGH}	0.070	0.875	−0.04
	825	0.930	2200 ₃₀₀	CC _{HIGH}	−0.052	0.455	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
2-chlorofuran (EH)	300	0.525	9 ₃	CC _{LOW}	−0.346	1.167	−0.09
	325	0.569	34 ₄	CC _{LOW}	0.107	1.154	0.06
				CC _{MID}	0.159	—	−0.01
	350	0.613	79 ₅	CC _{MID}	−0.050	—	−0.01
				CC _{HIGH}	−0.296	0.857	−0.09
	375	0.657	180 ₁₀	CC _{HIGH}	0.093	0.867	0.05
	400	0.701	350 ₄₀	CC _{LOW}	0.060	1.250	0.05
				CC _{MID}	0.033	—	−0.02
				CC _{HIGH}	−0.056	0.875	−0.05
	450	0.788	970 ₄₀	CC _{LOW}	−0.155	1.963	−0.06
				CC _{MID}	−0.027	—	−0.02
				CC _{HIGH}	−0.089	1.667	−0.05
	500	0.876	2300 ₂₀₀	CC _{MID}	0.052	—	0.03
				CC _{HIGH}	0.048	0.800	−0.04
	530	0.928	3300 ₂₀₀	CC _{HIGH}	−0.079	0.509	−0.04
2-chlorothiophene (EH)	300	0.459	1.3 ₃	CC _{LOW}	−0.327	2.500	−0.09
	350	0.536	15 ₅	CC _{LOW}	0.112	1.143	0.06
				CC _{MID}	0.094	—	−0.01
	375	0.574	35 ₂	CC _{MID}	−0.052	—	−0.01
				CC _{HIGH}	−0.131	0.400	−0.09
	400	0.613	81 ₅	CC _{MID}	0.032	—	0.03
				CC _{HIGH}	0.098	0.875	−0.09
	450	0.689	290 ₃₀	CC _{LOW}	−0.148	1.222	−0.09
				CC _{MID}	−0.011	—	−0.01
	500	0.766	820 ₄₀	CC _{LOW}	0.067	1.200	0.05
				CC _{MID}	0.067	—	−0.02
	550	0.842	1700 ₁₀₀	CC _{MID}	−0.030	—	−0.02
				CC _{HIGH}	−0.121	0.818	−0.05
	600	0.919	3300 ₃₀₀	CC _{HIGH}	0.056	0.833	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
phenol (EH)	360	0.519	3.6 ₄	CC _{LOW}	−0.136	3.667	−0.09
	440	0.634	65 ₅	CC _{MID}	−0.027	—	−0.01
	480	0.692	197 ₁₄	CC _{MID}	0.033	—	0.03
				CC _{HIGH}	0.075	1.750	−0.09
	540	0.778	710 ₁₁₀	CC _{LOW}	0.056	1.444	0.05
				CC _{MID}	0.032	—	−0.02
				CC _{HIGH}	−0.073	1.222	−0.05
	600	0.865	1870 ₁₅₀	CC _{HIGH}	−0.057	0.800	−0.04
1,2-dihydroxybenzene (EH)	480	0.594	24.1 ₁₄	CC _{LOW}	−0.600	1.083	−0.09
	500	0.619	60 ₁₀	CC _{LOW}	0.307	0.733	0.06
				CC _{MID}	0.288	—	−0.01
	520	0.644	80 ₂₀	CC _{MID}	−0.177	—	−0.01
				CC _{HIGH}	−0.554	0.923	−0.09
	550	0.681	180 ₂₀	CC _{LOW}	−0.147	0.582	−0.09
				CC _{HIGH}	0.419	1.364	−0.09
	580	0.718	360 ₃₀	CC _{MID}	0.093	—	0.03
	640	0.792	920 ₅₀	CC _{LOW}	0.092	1.734	0.05
				CC _{HIGH}	−0.253	1.719	−0.05
1,3-dihydroxybenzene (EH)	700	0.866	2000 ₁₅₀	CC _{MID}	−0.034	—	−0.03
	740	0.916	3300 ₁₃₀	CC _{HIGH}	0.053	0.577	0.04
	480	0.577	16.0 ₁₁	CC _{LOW}	0.063	1.528	0.06
	520	0.625	53 ₅	CC _{MID}	−0.025	—	−0.01
	580	0.697	240 ₅₀	CC _{MID}	0.039	—	0.03
	640	0.769	700 ₄₀	CC _{LOW}	−0.200	1.188	−0.06
				CC _{HIGH}	−0.107	1.719	−0.05
	700	0.841	1690 ₉₀	CC _{MID}	0.091	—	0.03
	760	0.913	3000 ₁₅₀	CC _{HIGH}	−0.168	0.842	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
1,4-dihydroxybenzene (EH)	480	0.565	7 ₃	CC _{LOW}	−0.766	1.528	−0.09
	520	0.612	41 ₅	CC _{LOW}	0.263	1.115	0.06
				CC _{MID}	0.303	—	−0.01
	550	0.648	79 ₆	CC _{LOW}	−0.174	0.582	−0.09
				CC _{MID}	−0.124	—	−0.01
				CC _{HIGH}	−0.501	0.655	−0.09
	580	0.683	180 ₄₀	CC _{LOW}	0.061	1.207	0.06
				CC _{MID}	0.110	—	−0.01
				CC _{HIGH}	0.235	0.897	−0.09
	640	0.754	550 ₃₀	CC _{LOW}	−0.161	1.188	−0.06
				CC _{MID}	−0.028	—	−0.02
				CC _{HIGH}	−0.298	1.719	−0.05
	700	0.824	1460 ₁₂₀	CC _{MID}	0.074	—	0.03
				CC _{HIGH}	0.051	0.829	−0.05
	760	0.895	2900 ₂₀₀	CC _{HIGH}	−0.136	0.842	−0.04
benzonitrile (EH)	298	0.435	0.20 ₁	CC _{LOW}	1.694	1.458	0.06
	360	0.525	2.0 ₁	CC _{LOW}	−0.557	1.331	−0.09
				CC _{MID}	−0.689	—	−0.01
	420	0.612	31 ₁	CC _{LOW}	−0.312	1.286	−0.09
				CC _{MID}	0.239	—	−0.01
				CC _{HIGH}	1.162	0.686	−0.09
	480	0.700	160 ₁₀	CC _{LOW}	0.282	1.250	0.06
				CC _{MID}	0.136	—	−0.01
				CC _{HIGH}	−0.418	0.751	−0.09
	540	0.787	450 ₆₀	CC _{LOW}	−0.104	1.778	−0.06
				CC _{MID}	−0.125	—	−0.02
				CC _{HIGH}	−0.242	0.778	−0.05
	600	0.875	1290 ₁₂₀	CC _{MID}	0.037	—	0.03
				CC _{HIGH}	0.226	0.800	−0.04
	640	0.933	2200 ₁₃₀	CC _{HIGH}	−0.059	0.563	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
p-benzoquinone (EH)	395	0.568	17.2 ₉	CC _{LOW}	0.194	1.359	0.06
	430	0.618	53 ₂	CC _{LOW}	−0.224	0.872	−0.09
				CC _{MID}	−0.082	—	−0.01
	460	0.661	140 ₂₀	CC _{MID}	0.119	—	0.03
				CC _{HIGH}	0.143	0.736	−0.09
	500	0.718	330 ₃₀	CC _{LOW}	0.093	1.200	0.05
				CC _{HIGH}	−0.256	1.147	−0.05
	550	0.790	810 ₆₀	CC _{MID}	−0.042	—	−0.02
	600	0.862	1850 ₇₀	CC _{HIGH}	0.078	0.833	0.04
naphthalene (EH)	423	0.564	18.0 ₄	CC _{LOW}	0.136	1.236	0.06
	473	0.631	65 ₂	CC _{MID}	−0.061	—	−0.01
	523	0.697	205 ₁₂	CC _{HIGH}	0.110	0.809	0.05
	575	0.767	550 ₂₀	CC _{MID}	0.033	—	0.03
	630	0.840	1230 ₁₂₀	CC _{MID}	−0.025	—	−0.02
				CC _{HIGH}	−0.062	0.878	−0.05
	700	0.933	3000 ₂₀₀	CC _{HIGH}	0.050	1.045	0.04
anthracene (EH)	550	0.615	22 ₄	CC _{LOW}	−0.164	2.273	−0.09
	600	0.670	76 ₁₀	CC _{LOW}	0.115	1.083	0.06
				CC _{MID}	0.050	—	−0.01
	625	0.698	122 ₁₅	CC _{MID}	−0.055	—	−0.01
	650	0.726	210 ₂₀	CC _{LOW}	−0.101	1.154	−0.06
				CC _{MID}	0.051	—	−0.02
				CC _{HIGH}	0.106	0.923	−0.05
	700	0.782	480 ₇₀	CC _{LOW}	0.098	1.143	0.05
				CC _{MID}	0.047	—	−0.02
				CC _{HIGH}	−0.143	1.786	−0.05
	750	0.838	900 ₉₀	CC _{MID}	−0.046	—	−0.02
				CC _{HIGH}	−0.088	0.867	−0.05
	800	0.894	1700 ₃₀₀	CC _{HIGH}	0.086	0.875	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
phenanthrene (EH)	550	0.613	24 ₁	CC _{LOW}	−0.115	2.273	−0.09
	600	0.669	76 ₁₁	CC _{MID}	0.035	—	0.03
	625	0.697	120 ₁₀	CC _{MID}	−0.020	—	−0.01
	650	0.725	190 ₂₀	CC _{HIGH}	0.038	0.923	0.03
	750	0.836	840 ₆₀	CC _{LOW}	−0.102	2.200	−0.06
	800	0.892	1520 ₁₅₀	CC _{MID}	0.032	—	0.03
	825	0.920	1900 ₁₀₀	CC _{HIGH}	−0.046	0.455	−0.04
naphthalen-2-ol (EH)	500	0.588	15 ₃	CC _{LOW}	−0.394	2.300	−0.09
	550	0.647	72 ₁₃	CC _{MID}	0.119	—	0.03
	575	0.676	120 ₃₀	CC _{MID}	−0.022	—	−0.01
				CC _{HIGH}	−0.171	0.435	−0.09
	600	0.706	200 ₃₀	CC _{HIGH}	0.043	0.917	0.03
	650	0.765	510 ₃₀	CC _{LOW}	−0.181	1.154	−0.06
				CC _{HIGH}	0.032	1.769	−0.05
	700	0.824	1090 ₁₅₀	CC _{LOW}	0.226	2.214	0.05
				CC _{MID}	0.084	—	−0.02
	750	0.882	1800 ₃₀₀	CC _{MID}	−0.070	—	−0.03
				CC _{HIGH}	−0.157	0.867	−0.04
	775	0.912	2500 ₃₀₀	CC _{HIGH}	0.102	0.452	0.04
naphthalene-2-carbonitrile (EH)	525	0.623	22 ₅	CC _{LOW}	0.149	0.921	0.06
	550	0.652	40 ₁₀	CC _{LOW}	−0.340	1.636	−0.09
				CC _{MID}	−0.077	—	−0.01
	580	0.688	90 ₁₀	CC _{LOW}	0.138	0.448	0.06
				CC _{MID}	0.129	—	−0.01
				CC _{HIGH}	0.162	1.086	−0.09
	600	0.712	120 ₁₀	CC _{MID}	−0.095	—	−0.02
				CC _{HIGH}	−0.208	0.611	−0.05
	650	0.771	310 ₃₀	CC _{HIGH}	0.307	2.231	0.03

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
quinoline (EH)	300	0.373	5.1 ₁	CC _{LOW}	2.988	3.000	0.06
	400	0.497	4.3 ₂	CC _{LOW}	0.794	1.250	0.06
				CC _{MID}	−0.747	—	−0.01
	450	0.559	11 ₁	CC _{LOW}	0.113	1.222	0.06
				CC _{MID}	−0.353	—	−0.01
				CC _{HIGH}	0.996	0.333	−0.09
	500	0.621	44 ₆	CC _{MID}	−0.051	—	−0.01
				CC _{HIGH}	0.635	0.800	−0.09
	550	0.683	150 ₄₀	CC _{MID}	0.036	—	0.03
				CC _{HIGH}	0.092	0.818	−0.09
	600	0.745	390 ₃₀	CC _{LOW}	−0.118	1.167	−0.06
				CC _{MID}	−0.021	—	−0.02
				CC _{HIGH}	−0.067	0.833	−0.05
	650	0.807	910 ₅₀	CC _{LOW}	0.085	1.423	0.05
				CC _{MID}	0.055	—	−0.02
				CC _{HIGH}	0.039	0.846	−0.05
	700	0.870	1700 ₂₀₀	CC _{MID}	−0.035	—	−0.03
				CC _{HIGH}	−0.101	0.857	−0.04
	740	0.919	2800 ₃₀₀	CC _{HIGH}	0.060	0.703	0.04
indole (EH)	400	0.522	2.2 ₂	CC _{LOW}	−0.422	1.250	−0.09
	450	0.587	22 ₃	CC _{LOW}	−0.225	1.222	−0.09
				CC _{MID}	0.188	—	−0.01
	500	0.652	99 ₉	CC _{LOW}	0.061	1.200	0.06
				CC _{MID}	0.101	—	−0.01
				CC _{HIGH}	−0.338	0.800	−0.09
	550	0.717	282 ₅	CC _{LOW}	0.052	1.182	0.05
				CC _{MID}	−0.028	—	−0.02
				CC _{HIGH}	−0.184	0.818	−0.05
	600	0.782	710 ₄₀	CC _{LOW}	−0.229	1.167	−0.06
				CC _{MID}	−0.024	—	−0.02
				CC _{HIGH}	0.051	0.833	−0.05
	650	0.847	1620 ₉₀	CC _{MID}	0.106	—	0.03
				CC _{HIGH}	0.044	0.846	−0.05
	700	0.913	2700 ₂₀₀	CC _{HIGH}	−0.196	0.857	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
isoindole (EH)	400	0.485	1.5 ₂	CC _{LOW}	−1.253	1.250	−0.09
	450	0.546	16 ₃	CC _{LOW}	0.755	1.222	0.06
				CC _{MID}	0.557	—	−0.01
	500	0.607	39 ₆	CC _{LOW}	−0.401	1.200	−0.09
				CC _{MID}	−0.340	—	−0.01
				CC _{HIGH}	−1.003	0.800	−0.09
	550	0.667	150 ₁₀	CC _{LOW}	0.213	1.182	0.06
				CC _{MID}	0.182	—	−0.01
				CC _{HIGH}	0.618	0.818	−0.09
	600	0.728	330 ₅₀	CC _{LOW}	−0.117	1.167	−0.06
				CC _{MID}	−0.098	—	−0.02
				CC _{HIGH}	−0.334	0.833	−0.05
	650	0.789	770 ₆₀	CC _{MID}	0.054	—	0.03
				CC _{HIGH}	0.180	0.846	−0.05
	700	0.850	1440 ₇₀	CC _{HIGH}	−0.100	0.857	−0.05
benzimidazole (EH)	500	0.531	2.3 ₁₃	CC _{LOW}	−1.391	1.200	−0.09
	550	0.584	31 ₃	CC _{LOW}	0.060	1.182	0.06
				CC _{MID}	0.632	—	−0.01
	600	0.638	85 ₉	CC _{LOW}	0.084	1.167	0.06
				CC _{MID}	−0.028	—	−0.01
				CC _{HIGH}	−1.159	0.833	−0.09
	650	0.691	210 ₃₀	CC _{LOW}	−0.146	1.154	−0.09
				CC _{MID}	−0.039	—	−0.01
				CC _{HIGH}	0.051	0.846	−0.09
	700	0.744	490 ₆₀	CC _{LOW}	0.246	1.143	0.05
				CC _{MID}	0.068	—	−0.02
				CC _{HIGH}	0.072	0.857	−0.05
	750	0.797	900 ₉₀	CC _{LOW}	−0.230	1.133	−0.06
				CC _{MID}	−0.115	—	−0.02
				CC _{HIGH}	−0.126	0.867	−0.05
	800	0.850	1900 ₂₀₀	CC _{MID}	0.108	—	0.03
				CC _{HIGH}	0.215	0.875	−0.04
	850	0.903	3000 ₃₀₀	CC _{HIGH}	−0.203	0.882	−0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
indazole (EH)	450	0.521	4.9 ₆	CC _{LOW}	0.571	1.222	0.06
	500	0.579	23 ₇	CC _{LOW}	−0.900	1.200	−0.09
				CC _{MID}	−0.257	—	−0.01
	550	0.637	130 ₂₀	CC _{LOW}	0.315	1.182	0.06
				CC _{MID}	0.409	—	−0.01
				CC _{HIGH}	0.467	0.818	−0.09
	600	0.695	260 ₃₀	CC _{LOW}	−0.093	1.167	−0.09
				CC _{MID}	−0.144	—	−0.01
				CC _{HIGH}	−0.750	0.833	−0.09
	650	0.753	610 ₇₀	CC _{MID}	0.043	—	0.03
				CC _{HIGH}	0.266	0.846	−0.05
	700	0.811	1170 ₁₁₀	CC _{LOW}	0.062	1.143	0.05
				CC _{HIGH}	−0.080	0.857	−0.05
	800	0.927	3700 ₃₀₀	CC _{HIGH}	0.055	0.875	0.04
purine (EH)	500	0.525	4 ₁	CC _{LOW}	0.061	1.200	0.06
	550	0.578	21 ₇	CC _{LOW}	−0.463	1.182	−0.09
				CC _{MID}	−0.028	—	−0.01
	600	0.630	88 ₁₃	CC _{LOW}	0.248	1.167	0.06
				CC _{MID}	0.212	—	−0.01
				CC _{HIGH}	0.051	0.833	−0.09
	650	0.683	200 ₃₀	CC _{MID}	−0.114	—	−0.01
				CC _{HIGH}	−0.391	0.846	−0.09
	700	0.735	500 ₅₀	CC _{LOW}	−0.147	1.143	−0.06
				CC _{HIGH}	0.213	0.857	−0.05
	750	0.788	1070 ₅₀	CC _{LOW}	0.052	0.778	0.05
				CC _{MID}	0.069	—	−0.02
	800	0.840	1830 ₉₀	CC _{MID}	−0.029	—	−0.02
				CC _{HIGH}	−0.129	0.875	−0.05
	875	0.919	3900 ₃₀₀	CC _{HIGH}	0.067	1.286	0.04

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
benzo[b]thiophene (EH)	348	0.452	14.4 ₁	CC _{LOW}	1.606	0.960	0.06
	398	0.516	10.6 ₁₃	CC _{LOW}	1.121	1.672	0.06
				CC _{MID}	−0.819	—	−0.01
	468	0.607	41 ₄	CC _{LOW}	−0.178	1.299	−0.09
				CC _{MID}	−0.420	—	−0.01
				CC _{HIGH}	1.672	1.041	−0.09
	523	0.679	180 ₂₀	CC _{MID}	0.077	—	0.03
				CC _{HIGH}	0.670	0.598	−0.09
	575	0.746	490 ₄₀	CC _{LOW}	0.071	2.250	0.05
				CC _{MID}	0.039	—	−0.02
				CC _{HIGH}	−0.137	0.770	−0.05
	650	0.843	1440 ₁₁₀	CC _{MID}	−0.022	—	−0.02
				CC _{HIGH}	−0.084	1.161	−0.05
benzo[c]thiophene (EH)	348	0.453	11.1 ₁	CC _{LOW}	2.437	0.960	0.06
	398	0.518	7.2 ₇	CC _{LOW}	0.065	1.672	0.06
				CC _{MID}	−1.243	—	−0.01
	468	0.610	58 ₁₅	CC _{MID}	−0.024	—	−0.01
				CC _{HIGH}	2.537	1.041	−0.09
benzoxazole (EH)	400	0.541	8.7 ₃	CC _{LOW}	0.192	1.250	0.06
	450	0.608	43 ₇	CC _{MID}	−0.085	—	−0.01
	500	0.676	180 ₁₀	CC _{LOW}	−0.181	1.200	−0.09
				CC _{HIGH}	0.153	0.800	−0.09
	550	0.743	550 ₃₀	CC _{MID}	0.082	—	0.03
				CC _{HIGH}	−0.054	0.818	−0.05
	600	0.811	1200 ₁₀₀	CC _{HIGH}	−0.151	0.833	−0.05

Table S5: (continued)

Molecule	T [K]	T_r	P_{sim} [kPa]	Test	$\ln(P_{\text{sim}}/P_{\text{CC}})$	ER	NB
benzisoazole (EH)	400	0.489	1.1 ₁	CC _{LOW}	−0.598	1.250	−0.09
	450	0.550	11 ₃	CC _{MID}	0.266	—	0.03
	500	0.611	43 ₆	CC _{LOW}	0.384	1.200	0.06
				CC _{HIGH}	−0.479	0.800	−0.09
	550	0.672	130 ₂₀	CC _{LOW}	−0.518	1.182	−0.09
				CC _{MID}	−0.174	—	−0.01
	600	0.733	450 ₅₀	CC _{LOW}	0.210	1.167	0.05
				CC _{MID}	0.238	—	−0.02
				CC _{HIGH}	0.320	0.833	−0.05
	650	0.795	830 ₅₀	CC _{MID}	−0.097	—	−0.02
				CC _{HIGH}	−0.439	0.846	−0.05
	700	0.856	1680 ₅₀	CC _{HIGH}	0.180	0.857	0.04
benzothiazole (EH)	400	0.508	2.8 ₃	CC _{LOW}	0.211	1.250	0.06
	450	0.572	15 ₃	CC _{LOW}	0.174	1.222	0.06
				CC _{MID}	−0.094	—	−0.01
	500	0.635	68 ₅	CC _{LOW}	−0.461	1.200	−0.09
				CC _{MID}	−0.078	—	−0.01
				CC _{HIGH}	0.169	0.800	−0.09
	550	0.699	270 ₂₀	CC _{LOW}	0.065	1.182	0.06
				CC _{MID}	0.210	—	−0.01
				CC _{HIGH}	0.142	0.818	−0.09
	600	0.762	580 ₂₀	CC _{LOW}	0.087	1.167	0.05
				CC _{MID}	−0.030	—	−0.02
				CC _{HIGH}	−0.384	0.833	−0.05
	650	0.826	1170 ₄₀	CC _{MID}	−0.040	—	−0.02
				CC _{HIGH}	0.055	0.846	−0.05
	700	0.889	2300 ₂₀₀	CC _{HIGH}	0.074	0.857	0.04

Table S6: All combined $\ln(Z_{\text{sim}}/Z_{\text{exd}})$ and $\ln(P_{\text{sim}}/P_{\text{CC}})$ Outliers in the TraPPE Development Data.

Molecule	T [K]	T_r	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	$\ln(P_{\text{sim}}/P_{\text{CC}})$
ethene (UA)	184	0.650	-0.037	-0.109
1,5-hexadiene (UA)	324	0.653	0.031	0.141
pentan-1-ol (UA)	400	0.691	0.039	-0.105
octan-1-ol (UA)	300	0.477	0.149	-0.184
	350	0.556	0.064	-0.222
	450	0.715	0.055	-0.144
	500	0.795	0.073	-0.114
propan-2-ol (UA)	300	0.598	-0.348	-0.974
	400	0.797	-0.072	-0.096
butan-2-ol (UA)	350	0.668	0.055	-0.408
2-methylpropan-2-ol (UA)	350	0.694	0.035	-0.537
1,3-propanediol (UA)	450	0.620	0.036	-0.541
ethanal (UA)	260	0.561	-0.364	-0.288
2-octanone (UA)	360	0.562	-0.540	-0.383
	400	0.625	-0.113	-0.105
ethylbenzene (UA)	397	0.640	-0.030	-0.288
isopropylbenzene (UA)	450	0.700	0.038	-0.132
o-xylene (UA)	450	0.711	0.055	0.065
m-xylene (UA)	451	0.719	0.057	-0.149
p-xylene (UA)	397	0.643	-0.157	-0.135
acetonitrile (UA)	398	0.725	-0.064	-0.068
nitrobenzene (UA)	298	0.399	-3.835	0.060
	375	0.502	-0.534	0.142
methanethiol (UA)	220	0.463	-0.419	-0.271
ethanethiol (UA)	240	0.478	0.106	0.072

Table S6: (continued)

Molecule	T [K]	T_r	$\ln (Z_{\text{sim}}/Z_{\text{exd}})$	$\ln (P_{\text{sim}}/P_{\text{CC}})$
pentanethiol (UA)	300	0.498	1.169	0.164
	320	0.532	0.618	0.093
2-butanethiol (UA)	280	0.503	-0.349	-0.149
2-methyl-1-propanethiol (UA)	300	0.539	-0.807	-0.335
	320	0.575	-0.299	-0.120
dimethyl disulfide (UA)	300	0.495	-0.419	-0.269
thiophene (UA)	293	0.484	-0.048	-0.474
pyrimidine (UA)	298	0.489	-0.260	0.288
	398	0.652	-0.072	-0.155
2-ethylhexyl acrylate (UA)	550	0.795	0.070	0.150
TIP4P-pol1 (pol)	373	0.643	0.051	0.073
n-dodecane (EH)	450	0.675	0.060	0.101
	500	0.750	0.068	0.156
dimethylamine (EH)	250	0.564	0.095	-0.165
	280	0.632	0.047	-0.093
ethylamine (EH)	293	0.648	0.050	-0.639
acetonitrile (EH)	398	0.728	-0.068	0.157
nitromethane (EH)	375	0.637	-0.038	-0.316
nitroethane (EH)	323	0.555	0.059	-0.094
pyridine (EH)	298	0.482	-0.092	-0.317
pyrimidine (EH)	298	0.472	-0.031	-1.528
pyrazine (EH)	348	0.565	0.116	-0.407
pyridazine (EH)	298	0.392	0.480	2.382
	348	0.457	-0.235	0.551
	398	0.523	-0.141	-0.489
	468	0.615	0.036	-0.177
furan (EH)	250	0.506	0.049	-0.413

Table S6: (continued)

Molecule	T [K]	T_r	$\ln (Z_{\text{sim}}/Z_{\text{exd}})$	$\ln (P_{\text{sim}}/P_{\text{CC}})$
pyrrole (EH)	300	0.461	-0.372	-0.783
	370	0.568	-0.041	0.109
thiazole (EH)	350	0.550	0.034	-0.525
	450	0.708	0.066	0.075
oxazole (EH)	300	0.557	-0.150	0.319
imidazole (EH)	400	0.489	0.148	0.121
pyrazole (EH)	350	0.475	0.092	-0.522
	400	0.543	-0.037	0.265
1,2,3-trichlorobenzene (EH)	400	0.532	-0.194	-0.535
hexachlorobenzene (EH)	575	0.648	-0.064	-0.273
1,2-dihydroxybenzene (EH)	500	0.619	0.092	0.307
1,4-dihydroxybenzene (EH)	580	0.683	0.048	0.061
benzonitrile (EH)	298	0.435	0.173	1.694
	540	0.787	-0.143	-0.104
p-benzoquinone (EH)	395	0.568	-0.043	0.194
	430	0.618	0.035	-0.224
anthracene (EH)	550	0.615	-0.458	-0.164
	600	0.670	-0.466	0.115
	650	0.726	-0.432	-0.101
	700	0.782	-0.453	0.098
phenanthrene (EH)	550	0.613	-0.467	-0.115
	750	0.836	-0.466	-0.102
naphthalen-2-ol (EH)	500	0.588	-0.408	-0.394
	650	0.765	-0.414	-0.181
	700	0.824	-0.448	0.226
naphthalene-2-carbonitrile (EH)	525	0.623	-0.405	0.149
	550	0.652	-0.445	-0.340
	580	0.688	-0.336	0.138

Table S6: (continued)

Molecule	T [K]	T_r	$\ln(Z_{\text{sim}}/Z_{\text{exd}})$	$\ln(P_{\text{sim}}/P_{\text{CC}})$
quinoline (EH)	300	0.373	-0.037	2.988
indole (EH)	400	0.522	0.121	-0.422
	450	0.587	0.079	-0.225
	500	0.652	0.121	0.061
	550	0.717	0.125	0.052
	600	0.782	0.113	-0.229
isoindole (EH)	400	0.485	-0.110	-1.253
	450	0.546	-0.102	0.755
	500	0.607	-0.059	-0.401
	550	0.667	0.125	0.213
	600	0.728	-0.069	-0.117
benzimidazole (EH)	500	0.531	0.092	-1.391
	700	0.744	0.072	0.246
indazole (EH)	500	0.579	0.105	-0.900
	550	0.637	0.190	0.315
purine (EH)	500	0.525	0.054	0.061
	550	0.578	0.117	-0.463
	600	0.630	0.148	0.248
benzisoxazole (EH)	600	0.733	0.179	0.210
benzothiazole (EH)	450	0.572	0.039	0.174
	550	0.699	0.156	0.065