

High-Throughput Molecular Dynamics Simulations and Validation of Thermophysical Properties of Polymers for Various Applications

*Mohammad Atif Faiz Afzal,^{*a} Andrea R. Browning,^a Alexander Goldberg,^b Mathew D. Halls,^b Jacob L. Gavartin,^c Tsuguo Morisato,^d Thomas F. Hughes,^e David J. Giesen,^e Joseph E. Goose^e*

^a Schrödinger, Inc., Portland, OR 97204, United States

^b Schrödinger, Inc., San Diego, CA 92121, United States

^c Schrödinger, Inc., Cambridge, Cambridgeshire CB1 2JD, UK

^d Schrödinger, Inc., K.K., Chiyoda-ku, Tokyo 100-0005, Japan

^e Schrödinger, Inc., New York, NY 10036, United States

* atif.afzal@schrodinger.com

Keywords: Molecular Dynamics (MD), graphics-processing-unit (GPU), polymers, glass transition temperature (T_g), melting temperature (T_m)

ABSTRACT: Recent advances in graphics-processing-unit (GPU) hardware and improved efficiencies of atomistic simulation programs allow the screening of a large number of polymers to predict properties that require running and analyzing long Molecular Dynamics (MD) trajectories. This paper outlines a MD simulation workflow based on GPU MD simulation and the refined Optimized Potentials for Liquids Simulation (OPLS) OPLS3e force field to calculate glass transition temperatures (T_g) of 315 polymers for which Bicerano reported experimental values.¹ Applying the workflow across this large set of polymers allowed for a comprehensive evaluation of the protocol performance and helps in understanding its merits and limitations. We observed a consistent trend between predicted T_g values and empirical observation across several subsets of polymers. Thus, the protocol established in this work is promising for exploring targeted chemical spaces and aid in the evaluation of polymers for various applications, including composites, coatings, electrical casings, etc. During the stepwise cooling simulation for the calculation of T_g , a subset of polymers clearly showed an ordered structure developing as the temperature decreased. Such polymers have a point of discontinuity on the specific volume vs. temperature plot, which we associated with the melting temperature (T_m). We demonstrate the distinction between crystallized and amorphous polymers by examining polyethylene. Linear polyethylene shows a discontinuity in the specific volume vs. temperature plot, but we do not observe the discontinuity for branched polyethylene simulations.

1. INTRODUCTION

Glass transition temperature (T_g) is one of the most essential characteristics of polymer materials defining their manufacturing processes and applicability range. In general terms, T_g refers to a temperature at which mechanical characteristics of the material undergo rapid changes upon cooling from the rubbery to a glassy state. This change in mechanical characteristics is in turn related to changes in characteristic relaxation times of the processes that can be specific to the materials chemistry. Depending on the material, T_g may exhibit various degrees of dependence on molecular weight and thermal history.²⁻⁴ Experimentally, T_g is measured by variations of three main techniques: Thermomechanical analysis (TMA),⁵ Dynamic Mechanical Analysis (DMA),⁶ Differential Scanning Calorimetry (DSC)⁷, and Broadband Dielectric Spectroscopy⁸. T_g measured by different methods may also display substantial variability.⁹ Computational methods have also been developed that allow calculating T_g . Bicerano¹ developed a correlation method that predicts T_g based on the molecular structure. According to Bicerano, structural parameters such as chain stiffness and cohesive forces provide reliable descriptors for a glass transition temperature description. However, Bicerano analysis does not give a detailed equation of state (Pressure-Volume-Temperature, PVT) for the polymers over the wide temperature range, which is often required for thermophysical characterization of a polymer. More detailed temperature dependence of polymer properties can be analyzed through Molecular Dynamics (MD) simulation.¹⁰⁻²⁵ The accuracy of MD predicted T_g has been evaluated by numerous groups. It is typically undertaken for small sets of polymeric systems²⁶ of a specific type (e.g., polyethylenes²⁷ and ethylene oxides²⁸, polyamides¹¹, polyimides^{23, 29}). In the MD approach, temperature-dependent properties, such as density, linear expansion parameters, or stiffness constants, are calculated via equilibration of polymeric systems at a broad range of temperatures. This procedure is computationally intensive. To manage such intensive calculations, MD protocols, including structure generation, equilibration, and analysis, are typically tuned to predict properties of a specific class of polymers. However, advances in polymer system builders, GPU accelerated molecular dynamics, equilibration techniques, and MD analysis allow for modeling of larger systems at longer simulation times, which can make T_g calculations automatic for arbitrary polymeric systems.

In this paper, we set out to investigate the feasibility of a MD screening approach for T_g prediction. In particular, we ask if a generic good-for-all protocol can be developed to estimate the glass transition in broad classes of polymeric systems reliably. We evaluate T_g of 315 polymers included in the Bicerano Handbook.¹ This list consists of a variety of polymers, which can be inferred by the wide range of experimental T_g values (130-685 K). The molecular weight of the monomer in these polymers ranges from 28 to 743 g/mol, which further corroborates the diversity in the polymer list. We identify several subsets of polymers with similar chemistry and demonstrate that the developed T_g model is efficient in rank-ordering of such polymers. Thus, the developed protocol is promising in exploring targeted chemical spaces for identifying high-performance polymers for various applications such as composites, coatings, electrical casings, etc. In addition to the validation results, we discuss the merits and limitations of our approach.

First and second-order phase transitions change the properties of materials. Many polymers are semi-crystalline, and therefore both crystallization and glass phase transition are pertinent to them. Crystallization is considered a phase transition of the first order³⁰⁻³³ when properties of a material have a point of discontinuity as a function of temperature, while glass transition is often be viewed as a second-order transition. The glass transition point can be recognized by a change in the slope of a linear plot of a thermophysical property as a function of temperature or as a kinetic (dynamic) change. Differences of opinions exist regarding the exact nature of the glass transition phenomenon and whether it is an actual second-order transition as well as the mechanism of the transition.³⁴ Some authors³⁵ argue that other types of discontinuities are observed at glass transition, which may be associated with local impurities, defects, and other local imperfections. This study focuses on the validation of experimentally observed glass transitions, and when a considerable discrepancy is found, this may be due to such other transitions. If a kinetic approach is taken, the glass transition can be found by exploring the segmental relaxation (for example Yin et al³⁶). Though the segmental relaxation approach provides interesting benefits in terms of focus on direct behavior of the polymer, in this study, we will only use the thermophysical approach of examining the visible discontinuities based on a detailed temperature dependence analysis as an evaluation of this widely used method.

As amorphous and semi-crystalline polymers undergo a glass transition, they change from glassy hard, stiff and brittle state to a viscous rubbery state.^{37, 38} When they do, their properties such as the coefficient of thermal expansion, the heat capacity, viscosity, refractive index, and others dramatically change.^{38, 39} The change associated with the glass transition can be observed and quantified using the dependence of density (specific volume) on the temperature. When a transition occurs, all the properties mentioned above experience a shift in the temperature derivatives.³⁸ This shift is attributed to the change in molecular motion. Although glass transition is an intrinsic characteristic of amorphous polymers, semi-crystalline polymers go through this phase transition too. Generally, the semi-crystalline polymer would have distinct crystallization and glass transitions. Its crystalline part undergoes melting at the melting temperature, while the amorphous region experiences glass transition at T_g . However, T_g is also affected by the crystalline component.⁴⁰ Although the primary purpose of this study is to validate glass transition temperatures of a set of polymers, our investigation addresses this effect as well. It should be noted that crystalline polymers have melting temperature always higher than glass transition temperature (if the latter exists). Therefore, upon cooling, crystallization occurs earlier than glass transition. In our study, we make no preliminary assumptions about the crystallinity of the polymers. Some of the polymers in our study develop ordering, and this process can be interpreted as partial crystallization. When such ordering happens in a polymer, we do not observe a clear glass transition in the simulation.

Empirical detection of glass transitions is based on two main approaches.^{38, 39} One is associated with the bulk properties, such as refractive index, stiffness, and hardness, while the other with the extent of molecular motion. It should be noted that experimental T_g depends on the

time scale of the experiment, which is typically in the order of minutes to hours. Such cooling rates are out of the scope of MD simulation time. This validation study is based on the experimental values provided by Bicerano¹ taken from different sources where different experimental methods were used, so scatter across the polymer systems in T_g value related to the experimental technique may be present. A more detailed validation with analysis of the agreement with different experimental methods is out of the scope of this study.

T_g calculations using molecular-level simulations have been performed since the 1990s.¹⁰ These methods either rely on the accurate prediction of the volume versus temperature behavior of polymers in both high- and low-temperature regions or the fitting the segmental relaxation (for example Starr and Douglas⁴¹ and Soldera and Grohens⁴²). These methods have proven to be sensitive to the selection of fitting in addition to the accuracy of the force field. To simulate T_g for a wide array of polymers, care must be taken to ensure that the method of data analysis is flexible enough to handle T_g 's that vary over 100's of degrees. The glass transition can also be forecasted by exploring the relaxations of molecules and projecting using established relationships between relaxation and time such as Vogel-Fulcher-Tammann.^{43, 44} Though this approach of relating the molecular relaxation to the glass transition is not taken in this study, it provides an interesting area for future study.

In Sec. 2, we introduce the dataset and simulation software used (Sec. 2a), provide simulation details (Sec. 2b), uncertainty quantification in T_g calculation (Sec. 2c), automated workflow for T_g calculation (Sec. 2d), protocol for studying the effect of chain length and polymer crystallization (Sec. 2e), and details on the coefficient of thermal expansion calculation (Sec. 2f). In Sec. 3, we present and discuss results for the density calculation (Sec. 3a), T_g model validation (Sec. 3b), crystallization onset studies (Sec. 3c), and report the coefficient of thermal expansion of polymers (Sec. 3d). We will highlight the accuracy of the T_g model by comparing it with the experimental results. We summarized our findings in Sec. 4.

2. METHODS

a. Dataset and simulation software

We screen through 315 polymers that have experimentally known glass transition temperatures. Those experimental temperatures, together with the values obtained through correlation models, are published in the book of Bicerano¹ in table 6.2. To validate the thermophysical properties workflow, that was developed as a part of Schrödinger Materials Science Suite (MSS), we calculated T_g of all the polymers reported by Bicerano. We performed all the calculations within Schrodinger MSS, version 2019-3.^{45, 46}

b. Simulation protocol

We built polymer chains such that the total number of atoms in a single chain is approximately 2000 atoms. This consistency in the number of atoms was used to keep the

approximate size of the polymer chains consistent across all the polymer systems given that the T_g depends on the molecular weight.⁴⁷ The number of repeat units varies for each polymer depending on the number of atoms in the repeat unit. For example, the polyethylene chain consisted of 333 monomers. A size of 2000 atoms was selected to ensure that all polymers in the study were of sufficient size to entangle. Subsequently, we constructed simulation cells with six polymer chains resulting in a box with approximately 12000 atoms. Larger systems of 16000 to 20000 atoms were also tested but similar T_g values resulted and 12000 atoms was used as the production system size. The initial density of the system in the amorphous cell structure was 0.5 g/cm^3 . After creating the amorphous structure at 0.5 g/cm^3 density, we equilibrated the system at 1000 K. This equilibration procedure consists of 20 ps of Brownian Dynamics at NVT (Number of atoms (N), Volume (V) and Temperature (T) are conserved) ensemble at 10 K followed by 20 ps of MD at NVT ensemble and finally 100 ps MD at NPT (Number of atoms, Pressure, and Temperature are conserved) ensemble. We performed all MD simulations in the workflow on GPU as implemented in Desmond code.^{46, 48} The simulations use the OPLS3e force field for all the systems in this work.^{39, 49-53}

The protocol for T_g calculation mimics the process of cooling from high temperature while simultaneously tracking the density changes of the system. For the majority of the polymers, we started the simulations from 1000 K and performed 5 ns NPT MD simulations at every 20 K step until 100 K. After 5 ns simulation at every temperature step, if the standard deviation of density exceeds 5%, additional simulations in steps of 5 ns is performed. For the polymer systems, whose expected T_g is higher than 600 K, we selected the temperature interval 1100 K to 200 K. For all styrene containing polymers, the starting temperature was 1100 K and the step size was 10 K instead of 20 K. This change in protocol for this large subset of the polymers was used to improve the precision by decreasing the temperature step size and increasing the temperature range. We used the time step of 1 fs for temperatures above 700 K and 2 fs time step for temperatures below 700 K. We used the Nose-Hoover thermostat^{54, 55} and the Martina-Tobias-Klein barostat.⁵⁶

To calculate T_g from MD simulations, researchers have typically used the specific volume as a function of temperature ($v(T)$).^{10, 11, 13-26} In the majority of these approaches, a bi-linear fit is used to compute the T_g value. The volume versus temperature dependence has a linear relationship for two regions, one below the T_g , which corresponds to the glassy coefficient of thermal expansion (CTE) and the other above the T_g , which corresponds to rubbery CTE. The intersection of the bi-linear fit in those two regions corresponds to T_g . However, for most of the polymers, the data does not show a clear delineation between the bi-linear region, and therefore such a fitting procedure may not be sufficiently robust for automation. In MD simulations, the transition between glassy and rubbery regions is spread over a range of temperatures. Thus, an additional drawback of the bi-linear fit approach is that we may have to remove points in the transition region to get better linear fits, which further limits the use of such a method in automated screening studies. Alternate approach to computing the T_g is by performing a hyperbolic fit to the density vs. temperature data. The hyperbola fit method was investigated recently by Patrone *et al.*⁷ This method utilizes the full range of data to identify the transition point from high temperature to low temperature behavior in

such a way that there is no user input required to define the high and low temperature regions. In this method, T_g is defined as the intersection point of the high temperature and the low temperature asymptotes to the fitted hyperbola. The hyperbolic fit was shown to overcome the ambiguity of the bi-linear fits and allows the automated estimation of T_g values.¹² The thermophysical properties analysis module in MSS includes both linear and hyperbolic fitting. However, in this work, we adopt the hyperbolic fitting approach due to the limitations of the bi-linear fit described above.

c. Uncertainty Quantification

Extracting the T_g values by performing hyperbolic fit to density vs. temperature data also results in fitness or within simulation uncertainty. Additionally, as MD simulations are dependent on the initial configuration of the system, there is also uncertainty between simulations. Therefore, to obtain better statistics, it is typically recommended to perform simulations on multiple replicas. In this work, we generate ten simulation cells with different initial configurations (replicates) for each polymer. We then execute the T_g calculation protocol on each of these replicas and compute the T_g value by performing the hyperbolic fit. We calculate the aggregate T_g and standard deviation of the polymer by taking the weighted average of T_g values from all the ten replicas, where weights are the inverse of the within simulation uncertainty as described in Patrone.¹²

d. High-throughput screening

The computation of the T_g for each system involved a large number of MD simulations. For example, a typical T_g calculation will require 200-400 ns total simulation time. In this work, which includes T_g calculations of ten replicas for each of 315 polymer systems, we performed a total of over 1 million ns of MD simulation. We developed an automated framework to perform these large scale simulations, as shown in Figure S1. We first converted the polymer names from the Bicerano handbook to machine-readable SMILES (Simplified Molecular Input Line Entry System) strings.⁵⁷ The SMILES strings include the monomer structure along with the information of the head and tail of the monomer. The automated process for computing the T_g values include: 1) extraction of SMILES of the monomer, 2) creation of polymer chain for all 315 polymers, 3) generation of ten replicas for each polymer, 4) running the T_g protocol for each replicate, 5) performing hyperbolic fit to compute the T_g value, 6) collection of T_g values from different replicas and performing the uncertainty quantification, and 7) resubmission of T_g jobs, with a larger temperature window in order to capture deviations caused by insufficient sampling of high and low temperature region, if the standard deviation between replicas is more than 50 K. Additionally, a guess value for the T_g can also be provided to select a narrow temperature window in the hyperbola fit.

Due to the high computational costs associated with polymer MD simulations, there has been significantly less number of high-throughput MD screening efforts when compared to quantum mechanics (QM) based screening.⁵⁸⁻⁶¹ Simmons et al^{62, 63}, and others⁶⁴ performed high-throughput studies to compute the T_g of glass-forming liquids. Most of these MD based screening

studies for T_g calculations were limited to less than 100 polymers. However, with the advancement in hardware and efficient implementation of GPU accelerated MD code, we are now able to perform MD screening for several hundreds of polymers in just a few days. Using the automated T_g workflow in this work, we conducted more than 150,000 simulations (including replicates and individual temperature runs). An alternative to MD approach is to apply machine learning based models to accelerate the screening of polymers for T_g prediction.^{65, 66} Machine learning models can predict T_g in a fraction of a second; however, the accuracy of such models is highly dependent on the availability of data.⁶⁷ Since the data from experiment is limited, MD based models could provide additional data for building efficient machine learning models.

e. Effect of polymer chain length and crystallization

In our primary T_g protocol, we used polymer chains containing approximately 2000 atoms. When using such length chains in MD simulation, we generally do not observe the ordering of polymer chains. Therefore, we also develop simulation cells for all of the polymers in the study with a mixture of long and small polymer chains and cast these cells through our T_g protocol. A mixture of two different chain lengths is known to enhance crystallization for polymer materials.⁶⁸ Details of the simulations cell and T_g protocol for these systems are provided in the SI (Supporting Information). It should be noted that we only use a single replicate for these simulation cells as the primary purpose of this approach is to check for possible ordering during the cooling process. For example, such simulations would allow us to see the formation of ordered chains that correspond to the onset of crystallization. We report, in the main text, the polymers where we observe ordering.

f. Coefficient of thermal expansion

As the process of T_g computations involves density calculation at a range of temperatures, we can use the data to evaluate the CTE below the T_g (glassy CTE at 300 K) as well as above the T_g (rubbery CTE at 800 K). We computed the glassy CTE and rubbery CTE of all the 315 polymers by performing linear regression on the specific volume vs. density data from the ten lowest temperature points and ten highest temperature points, respectively. We do not calculate the glassy CTE of the polymers whose T_g value is less than 300 K. We compare the results for all the polymers and additionally report the correlations between the T_g and CTE properties.

3. RESULTS AND DISCUSSION

a. Density results

As the process of obtaining T_g includes the evaluation of density at various temperatures, we also obtain the density values at 300 K. We use these values to compare with the experimental density values to provide additional validation of our computational models. Among the 315 polymers investigated, we found the density of 150 polymers that are reported by Bicerano,¹ in table 3.5. MD predicted density of most of the polymers differs by less than 6% compared to

experimental values (correlation coefficient $R^2=0.95$). Figure 1 shows the comparison of calculated and experimental density values, where the colormap shows the experiment T_g values of individual points. The R^2 of 0.95 confirms a good agreement with the experimental density values. The benchmark comparison gives a mean absolute error (MAE) of 0.040 (3.48%), a root mean squared error (RMSE) of 0.052 (4.39%), and a maximum error (MaxE) of 0.205 or 17.86%, respectively, demonstrating that our modeling protocol is accurate. The average error (AE) is small with -0.022 (-2.18%), i.e., MD approach is not significantly biased towards systematic over- or under-prediction. The slope (1.004) and intercept (-0.027) of the linear fit line further corroborates this observation. The outliers with the prediction error of more than 10% include the multi-ringed polymer poly(quinoxaline-2,7-diylquinoxaline-7,2-diyl-1,4-phenylene). Poly(o-methyl styrene) also shows a deviation from experiment of greater than 10% despite other styrene polymers resulting in more accurate densities. Further investigation is needed to understand the source of this discrepancy. The color of each point in Figure 1 represents the glass transition temperature of that polymer. We notice that the outliers in this figure typically have higher T_g values. We split the figure into two plots (Figure S2a and Figure S2b): one with all polymers whose T_g values are less than 300K and the other with all polymers whose T_g values are higher than 300K. We observe that there is better agreement with experiment values for the polymers whose T_g values are less than 300K, i.e., the polymers that exhibit rubbery behavior at 300K.

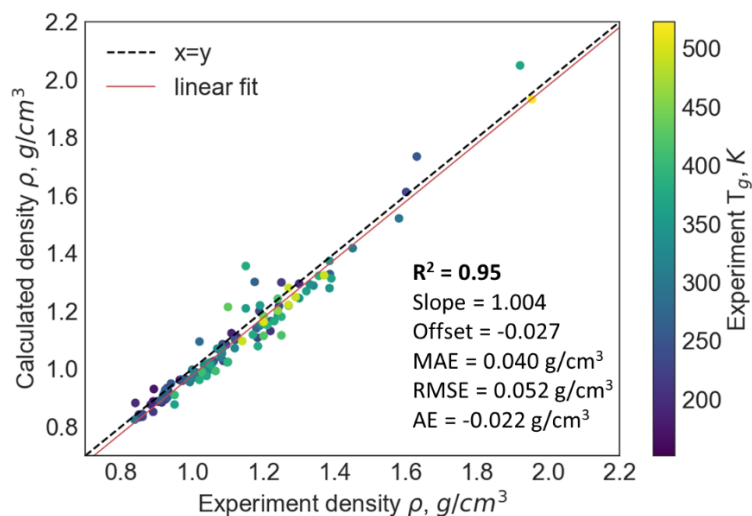


Figure 1. Predicted density values at $T=300$ K against experimental density reported in Bicerano¹ table 3.5.

b. T_g model validation

We extracted T_g for our systems by performing a hyperbola fit to the density against temperature plot, where the T_g value corresponds to the intersection point of high and low temperature asymptotes.¹² We show an example of the automated hyperbola in Figure 2 for polyoxyoctamethylene polymer. Based on the fitting, the T_g of this polymer is calculated as 238 K, which is overestimated compared to the experimental T_g (218 K) value for this polymer. It

should be noted that the experimental determination of the T_g values takes place within minutes or even hours. It is known,^{10, 69} that extending the time of the volumetric measurements results in a decrease of T_g values. Although the state-of-the-art MD computation methods allow running dynamics on GPU for microseconds,⁴⁸ the time scales are still significantly shorter than typical times in experiments. Thus, T_g values are expected to be overestimated compared to experiment by approximately 3 K/order of magnitude rate difference assuming Williams-Landel-Ferry (WLF) behavior and universal parameters for polymers.²⁰ The (WLF) parameters vary between polymers causing variation in the rate dependence for different polymers.⁷⁰ Establishing a direct correlation between the calculated and experimental methods can result in a correction that can be applied to simulated T_g values to improve accuracy.

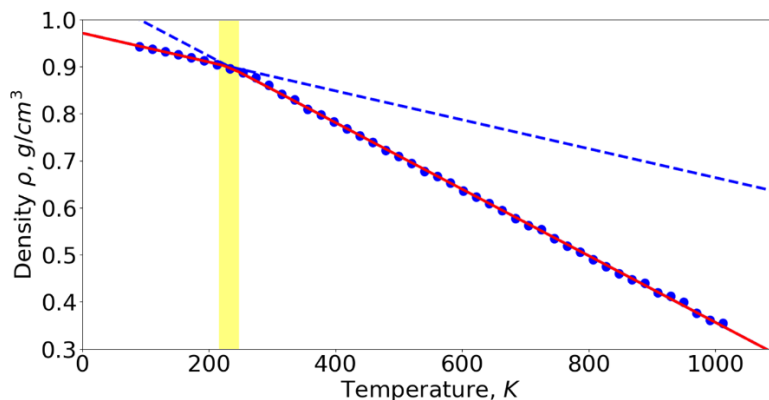


Figure 2. Hyperbola fit of density as a function of temperature to evaluate a glass transition temperature (T_g) of polyoxyoctamethylene. The red curve represents the hyperbolic curve, and blue dashed lines represent the asymptotes. The yellow region represents the portion of simulated density values that deviates 10% from the asymptotes.

We obtained T_g values for all the ten replicates for each polymer and subsequently computed the aggregate T_g of polymer using the method described in Sec. 2b. Computed T_g values along with the standard deviation for all 315 polymers are shown in Figure 3. We included the absolute values in the supporting information. As expected, for most of the polymers, observed T_g values are overestimated (average deviation is +79.1 K) compared to the experimental values. However, the R^2 value of 0.92 shows that there are clear trends with the experiments. The linear fit, with a slope of 1.30, suggests that the prediction overestimation increases with T_g . As the polymers in this set of 315 polymers are structurally diverse, the offset we obtain from this linear fit could be used as a calibration factor for other polymers as well. Based on the linear fit, the calibrated T_g is calculated as $0.77 * T_g(\text{calc.}) + 21.08 \text{ K}$. We show the calibrated T_g values of all the 315 polymers in Figure 4. The prediction errors for the calibrated T_g are reasonable, with a mean absolute error (MAE) of 27.5 K (7.98%) and a root mean squared error (RMSE) of 36.18 (10.92%). The shaded area in Figure 4 indicates all the points that are within 10% error from the experimental values. There are a total of 272 polymers that have less than 10% error and 226 polymers with less than 5% error.

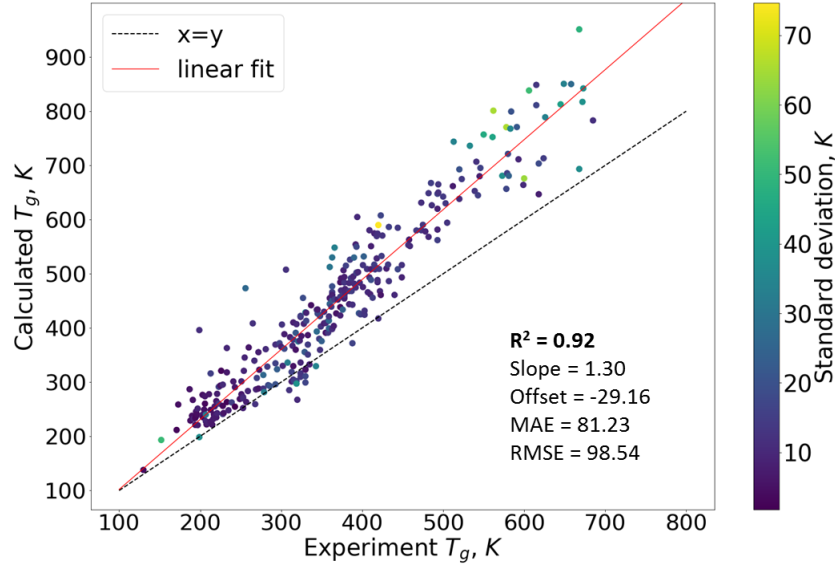


Figure 3. Predicted T_g values against experimental ones reported in Bicerano¹ Table 6.2. The color bar represents the standard deviation of prediction based on the uncertainty quantification.

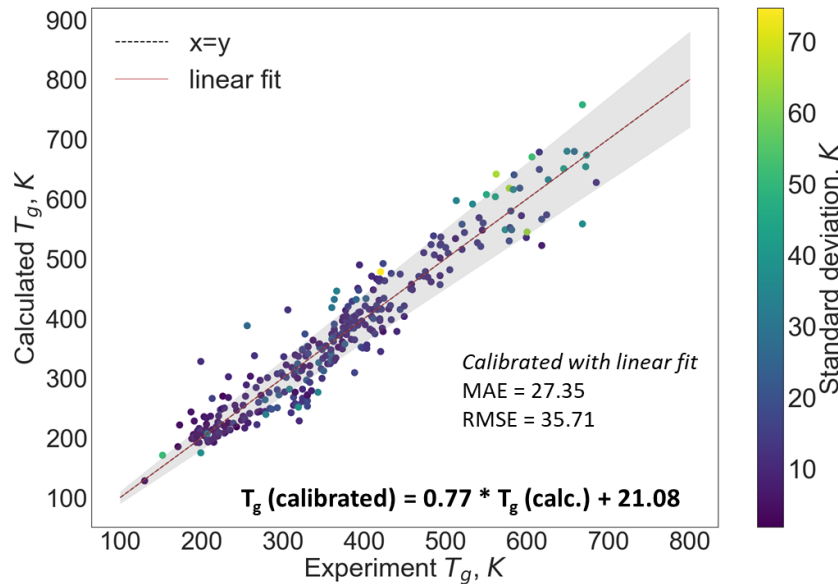


Figure 4. Calibrated T_g values against experimental ones reported in Bicerano¹ Table 6.2. The color bar represents the standard deviation of prediction based on the uncertainty quantification. The points in the shaded area within 10% error from the experiment value.

The color of the points in Figure 3 and Figure 4 show the standard deviation of T_g prediction, which is estimated using the uncertainty quantification as described in Sec. 2c. We observe larger deviations from the experimental values at high temperatures. This may be caused by the behavior of the force field as the simulation temperature deviates from the force field fitting temperature or due to the differences in the rate dependence of T_g between polymer systems. Deviation of the predicted T_g values from the experiment can be evaluated based on the

temperature difference $T_g(\text{predicted}) - T_g(\text{experiment})$ (Figure 5). Figure 5a shows the deviations of the calculated T_g from the experimental T_g , whereas Figure 5b shows the differences between calibrated T_g and experiment T_g . The entire distribution of calibrated T_g is close to the Gaussian curve, with the majority of the polymers showing less than 30 K deviation. We also observed that the deviation in predicted T_g was larger for some classes of polymers such as styrene containing polymers. In this study, we explored corrections that depend only on T_g and leave for future work the exploration of chemistry dependent corrections.

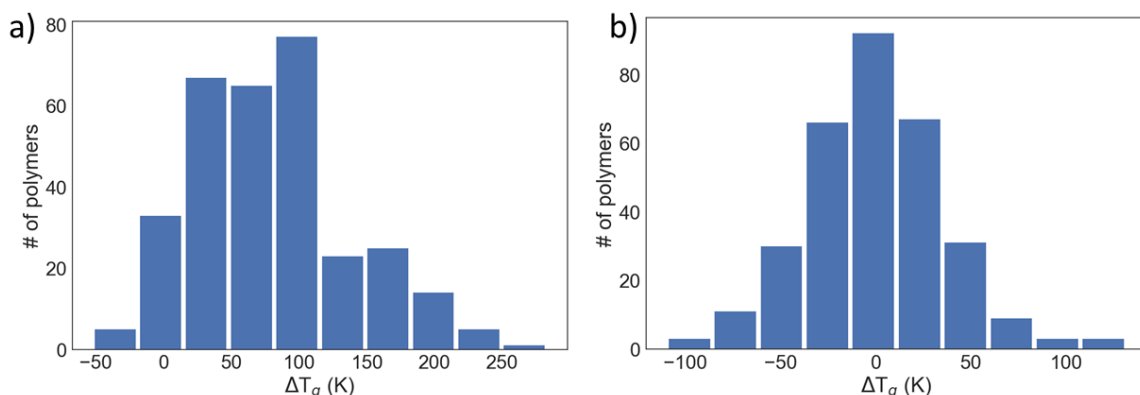


Figure 5. Histogram showing the deviations of the T_g from; a) the parity line, and b) the linear regression fit.

Although the MAE of 27.5 K may suggest that the T_g model is not performing very well, we would like to stress that the large errors may be primarily due to the diverse nature of the polymer set along with the inconsistencies in the experimental values. For example, a range of 30 K difference is observed in experimental measures of a single thermoset polymer system.⁷¹⁻⁷³ We believe that the model validation is more meaningful when we apply this to subsets of polymers with similar chemistries. Therefore, we selected specific types of polymers to further validate our model. The subsets include common types of polymers such as polyacrylates, polyamides, polyimides, and subsets with polymer names containing styrene, propylene, butadiene, phenyl, vinyl, and oxy. In addition to these, we perform analysis on all the polymers that contain only C and H, polymers containing at least one halogen, and polymers containing silicon. Figure 6 shows the comparison of computed and experiment values for all these subsets. We obtained the computed values using the calibration scheme presented in Figure 4, and considered only the points with computation uncertainty less than 10 K. For all the polymer subsets considered, except for polyimides and polyamides, the T_g protocol is performing well, with the majority of subsets showing an MAE of less than 20 K. For polyimides and polyamides, the model gives an MAE of 45.5 K and 32.0 K, respectively, suggesting that the current model may give large T_g prediction errors for this class of polymers. However, despite the large errors, the model is getting correct order of T_g values for these polymer systems, which can be confirmed from the plots as well as from the high R^2 value. In contrast, for styrene containing polymers, we see a low R^2 (0.4) value but the model performance is significantly better (MAE = 14.2 K). Polymers containing either oxygen, nitrogen, or silicon also show better model performance when compared to the performance on the complete set. This demonstrates that the current MD protocol is a good tool to correctly rank-order the T_g 's within the majority of the polymer subsets. Such a rank ordering method can aid in exploring new polymers with required T_g values in

targeted chemical spaces. Additionally, such analysis provides valuable information to further improve the protocol improvement for specific classes of polymers.

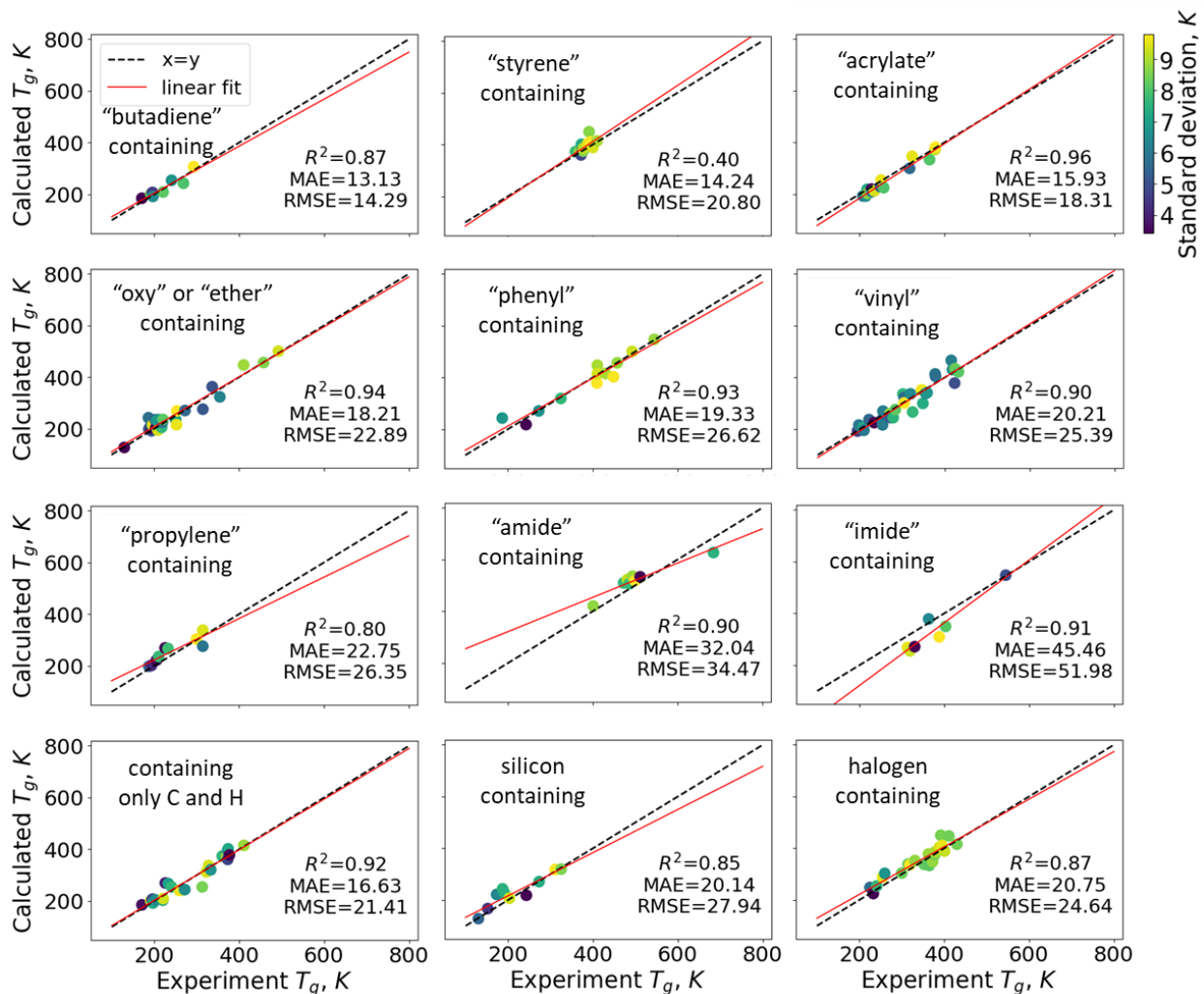


Figure 6. Comparison of computed and experimental values for subsets of polymers from the Bicerano dataset. The color bar represents the standard deviation of T_g prediction based on the uncertainty quantification.

c. Effect of chain length and onset of polymer crystallization

In MD simulations, where we use long polymer chains of single molecular weight, it is challenging to observe any effect of chain ordering during typical MD time-scales. To observe any chain ordering behavior, we developed simulation cells for all 315 polymers with a mixture of small and long chains, as described in SI. In some of these polymeric systems, the volume behavior reveals some phase changes that are not associated with the glass transition. The transitions are related to ordering that can be recognized as the breaking point in the specific volume vs. temperature plot. In the case of polyethylene, the onset of the crystallization temperature region can be localized on the $v(T)$ plot of the linear polyethylene (Figure 7). Since the T_g calculation

starts at a high temperature, the crystallization transition can be identified in the 320-360 K temperature range. According to the experiment,³⁹ the melting temperature is in the range of 380-420 K. Our calculation underestimates the onset of melting that is apparently due to the overcooling of the prepared system.

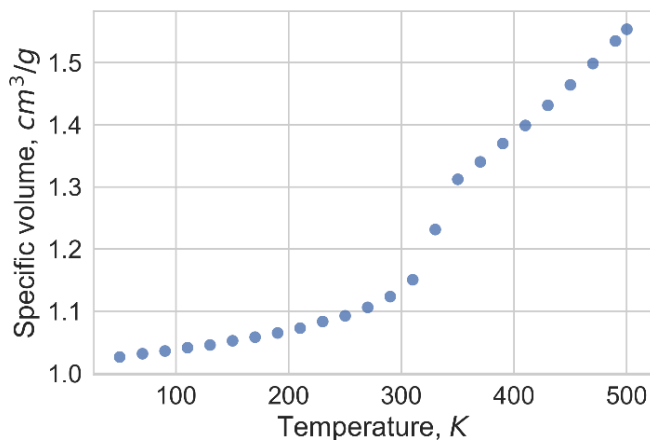


Figure 7. Specific volume as a function of temperature for linear polyethylene polymer. The transition is associated with the melting temperature (T_m).

Other polymers show an evident crystallization point in the $v(T)$ plot. One of the interesting examples is poly(vinyl alcohol) polymer. According to the experiment, its melting temperature is 473 K and its boiling point is 501 K.⁷⁴ According to MD calculations, at the temperature of 600 K, the polymer chains in the unit cell are forming into an ordered structure, where polymer chains are elongated (Figure 8a). As the temperature continues to decrease, the polymeric structure is arranged into a more regular shape that has some characteristics of an ordered structure. At the same time, it is not completely crystalline as the chains maintain some curvature, and dihedral angles orientations in the polymer chains are not aligned. Although this structure is not fully crystalline, the T_g of such a geometrical arrangement cannot be determined. The imperfect crystalline structure can be attributed to a short simulation time compared to the experimental time window. According to some publications,^{30, 31} the T_g of the semi-crystalline polymer should correspond to the T_g of its amorphous part. According to our fitting protocol, the T_g in such a case cannot be determined unambiguously (Figure 8b). The obtained imperfect crystalline structure contains too small an amorphous fraction to give an estimate of the T_g value.

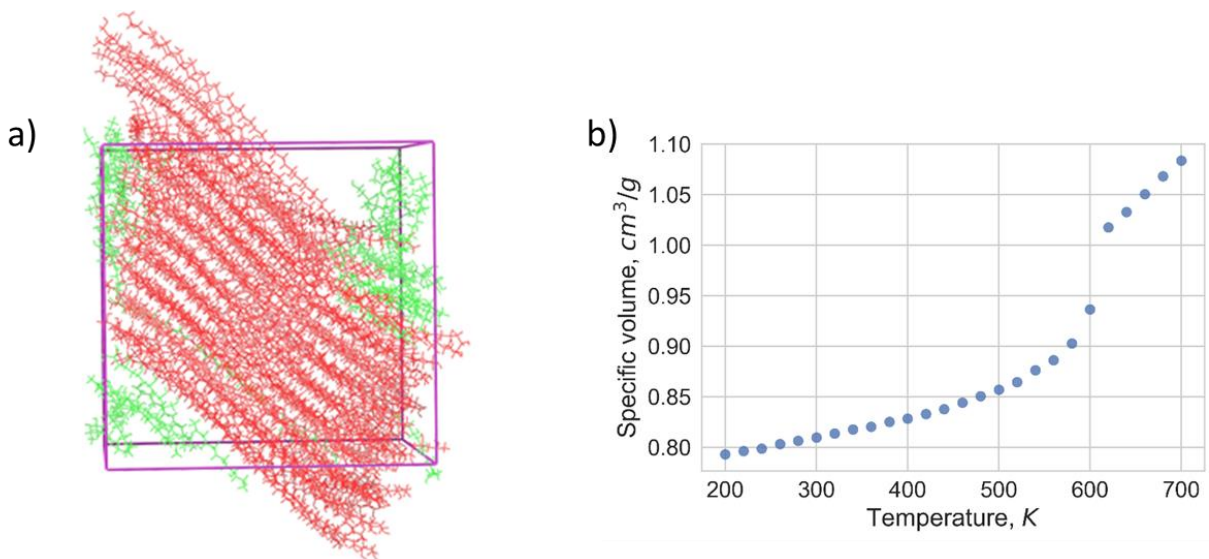


Figure 8. Snapshot of the polyvinyl alcohol polymer at 600 K (a) and specific volume as a function of temperature (b). The onset of crystallization can be identified in the plot.

Based on the observation of other transitions described above, we reviewed the results for the polymers in the test set where estimated T_g deviates from the experiment by more than 100 K. We confirm the chain ordering in the polymers by visual examination of their atomic structure at different temperatures and comparing it to the specific volume against temperature plots. We observed ordering in several polymers including (poly(vinyl acetate), linear polyethylene, poly(vinyl alcohol), poly(p-hydroxybenzoate), poly(p-phenylene terephthalamide), poly(vinylidene chloride), poly(neopentyl methacrylate), poly(3,3-dimethylbutyl methacrylate), poly(p-isopentoxy styrene), poly(bromo-p-xylylene), poly(2-methyl-5-t-butyl styrene), poly(a,a,a',a'-tetrafluoro-p-xylylene), poly(o-vinyl pyridine), poly(vinyl formal), poly[2,2-butane bis{4-(2-methylphenyl)}carbonate], poly(N-vinyl carbazole), poly [1,1 -cyclopentane bis(4-phenyl)carbonate], polyetherimides, poly(bisphenol-A terephthalate), Ultem, polyimides, and polyquinoline. Once they transition into an ordered structure, the transition to a glassy phase cannot be observed.

To avoid crystallization, some polymers such as polyethylene can be branched. Since branching prevents the polymer crystallization, the T_g can be evaluated more accurately in the branched polymer. As more branching points are included in building the polymer, the less crystallization temperature is pronounced. Figure 9 shows the specific volume variation for three different branched polyethylene: branched polyethylene with two branching generations and 50% probability (b2-pc50), branched polyethylene with two branching generations and 75% probability (b2-pc75), and branched polyethylene with three branching generations and 25% probability (b3-pc75). In all the branched polyethylene polymers, the melting temperature is not unambiguously observed. Thus, as we introduce branching, the melting transition becomes less pronounced. As the melting becomes less pronounced, the transitions associated with T_g are more distinct and therefore is straightforward to determine the T_g value using a hyperbolic fit.

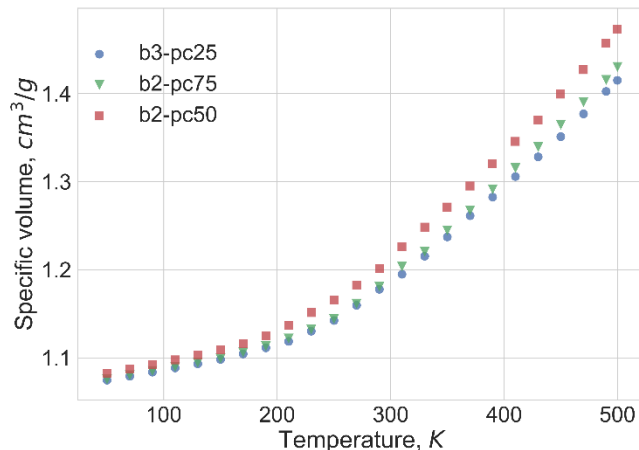


Figure 9. Specific volume as a function of temperature for three branched polyethylene polymers, polyethylene with two branching generations and 50% probability (b2-pc50), with two branching generations and 75% probability (b2-pc75), and with three branching generations and 25% probability (b3-pc75).

d. Coefficient of thermal expansion

Using the specific volume plots obtained from the T_g protocol, we computed the coefficient of thermal expansion (CTE) of the polymers at 298 K, which correspond to the glassy CTE for most of the polymers in our list as described in Sec. 2.f. We also computed the CTE at 700 K, which refers to the rubbery CTE. The distribution of calculated glassy CTE values, for the polymers with experimental T_g greater than 298 K, and the rubbery CTE values are shown in Figure 10a and Figure 10b, respectively. Glassy CTE for these polymers ranges from $100 \times 10^{-6} \text{ K}^{-1}$ to $500 \times 10^{-6} \text{ K}^{-1}$, with a median value of $200 \times 10^{-6} \text{ K}^{-1}$. We included the standard deviation of calculation, for both glass CTE and rubber CTE, in the color bar of Figure 10c and Figure 10d, respectively. In the case of glass CTE values, the standard deviation ranges from $2.7 \times 10^{-6} \text{ K}^{-1}$ to $19.5 \times 10^{-6} \text{ K}^{-1}$. We provide the calculated CTE values of the polymers in the SI. Figure 10c shows the correlation between the calculated glassy CTE and rubbery CTE values, which shows a linear trend. The standard deviation (color bar) in this plot refers to the deviation between the CTE values among the ten replicates. Figure 10d shows the correlation between the calculated glassy CTE values with the experimental T_g values. We observe that the CTE values decrease with T_g values, which is in agreement with the data reported in Bicerano figure 3.2.¹ We notice that the standard deviation is higher for the polymers whose T_g is close to 300 K. This is because the variation between specific volume and temperature is not linear near 300 K for such polymers.

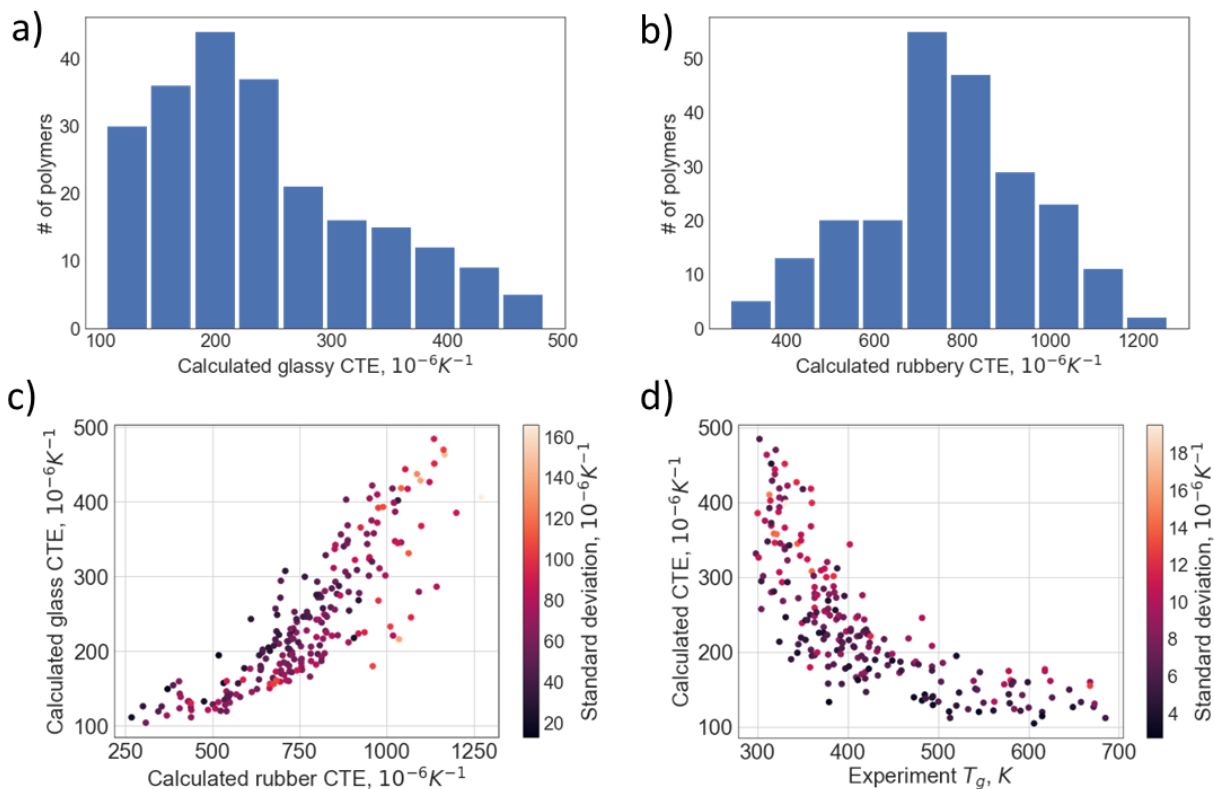


Figure 10: a) Distribution of glassy CTE values of the polymers; b) Distribution of rubbery CTE values of the polymers; c) Comparison of glassy CTE and rubber CTE values with the color bar showing the standard deviation of rubber CTE calculation; d) Variation of glassy CTE values with the experimental T_g (color bar shows the standard deviation of glass CTE calculation).

4. CONCLUSIONS

This study shows that with current state-of-the-art modeling software, together with recent advances in the GPU hardware, it is possible to systematically screen a large number of polymers to predict glass transition temperatures (T_g). For the first time, the accuracy of molecular dynamics simulations for T_g prediction was validated across an extensive range of polymers. These high-throughput MD simulations, based on GPU accelerated Desmond code, consumed more than 10 years of GPU time. In addition to validating against the complete dataset, we evaluated the performance of the T_g model on subsets of polymers with similar chemistries. The validation studies for the several subsets of polymer chemistries show a trend between computed and experimental values indicating that our model is useful in estimating the T_g and for the correct rank-ordering of polymers within the subsets. The ability to rank order will allow us to design high-performance polymers in targeted chemical space.

Polymers that have ordered structure were found to have a point of discontinuity on the specific volume vs. temperature plot. This discontinuity is a clear indication of the melting onset.

We confirmed the partial crystallization of those polymers by analyzing structures of those polymers below the transition temperature. We demonstrated the distinction between crystallized and non-crystallized polymers by studying polyethylene. Linear polyethylene experiences crystallization in the 320-360 K temperature range. As branching points are added to polyethylene the crystallization temperature is less pronounced. The specific volume vs. temperature dependence shows a break in those plots. Due to the semi-crystalline complexity of those polymers, we assume that those points can be associated with different other possible local phase transitions. In addition to T_g evaluation, we successfully computed the CTE of polymers and showed the correlation between CTE and T_g values.

We have demonstrated that T_g can be obtained in measurable agreement with the experiment within subsets of chemistries using our T_g protocol that provides efficient calculation times. The hyperbola fitting method, implemented in our work, allows us to perform automated screening of polymers by eliminating the influence of user-selected high and low temperature regions. The polymers that are included in this work comprise of diverse chemical structures which is further corroborated by the broad range of experiment T_g (ranging from 130 K to 685 K). If the expected T_g of a new polymer is within this range, we believe the protocol presented in this work would be a good starting point to compute T_g value. Additionally, the presented automated framework to compute the T_g of polymers would aid in accelerated property prediction of a large set of polymers. Thus, the framework is promising for the design and discovery of new polymers in targeted molecular space for various applications. In contrast to the correlation method used by Bicerano,¹ we not only compute the T_g values but also detect other phase transitions that occur in polymer systems. Our work is furthermore an example of the high-throughput study in thermodynamics of polymers that has become possible due to the recent advancement in computer hardware technology allowing us to run relatively long MD trajectories using state-of-the-art acceleration on GPU.

5. SUPPORTING INFORMATION

Electronic supporting information (SI) accompanies this paper and is available through the journal website. It provides the calculated T_g values of all 315 polymers along with the experimental T_g that were extracted from the Bicerano book. The SI also includes calculated density, experimental density, CTE, and monomer SMILES of the polymers. The head and tail of the polymer are represented as Ce and Th atoms in the SMILES of the monomer. We also give detailed information on statistical metrics used in this work. Additionally, we include the simulation protocol used to evaluate the onset of chain ordering in the polymer cells.

6. AUTHOR INFORMATION

Corresponding Author: Mohammad Atif Faiz Afzal

*E-mail: atif.afzal@schrodinger.com

REFERENCES

1. Bicerano, J., *Prediction of Polymer Properties*. Marcel Dekker Inc.: New York, 1996.
2. Zhang, J.; Liang, Y.; Yan, J.; Lou, J., Study of the molecular weight dependence of glass transition temperature for amorphous poly(l-lactide) by molecular dynamics simulation. *Polymer* **2007**, *48* (16), 4900-4905.
3. Dalnoki-Veress, K.; Forrest, J.; Murray, C.; Gigault, C.; Dutcher, J., Molecular weight dependence of reductions in the glass transition temperature of thin, freely standing polymer films. *Physical Review E* **2001**, *63* (3), 031801.
4. Keddie, J. L.; Jones, R. A.; Cory, R. A., Size-dependent depression of the glass transition temperature in polymer films. *EPL (Europhysics Letters)* **1994**, *27* (1), 59.
5. Schwartz, A., Glass transition temperatures of polymer materials, measured by thermomechanical analysis. **2005**, *13* (3), 489.
6. Pinheiro, A.; Mano, J. F., Study of the glass transition on viscous-forming and powder materials using dynamic mechanical analysis. *Polymer testing* **2009**, *28* (1), 89-95.
7. Richardson, M.-J.; Savill, N., Derivation of accurate glass transition temperatures by differential scanning calorimetry. *Polymer* **1975**, *16* (10), 753-757.
8. Kremer, F.; Schönhals, A., *Broadband dielectric spectroscopy*. Springer Science & Business Media: 2002.
9. Backfolk, K.; Holmes, R.; Ihalainen, P.; Sirviö, P.; Triantafillopoulos, N.; Peltonen, J., Determination of the glass transition temperature of latex films: Comparison of various methods. *Polymer Testing* **2007**, *26* (8), 1031-1040.
10. Han, J.; Gee, R. H.; Boyd, R. H., Glass transition temperatures of polymers from molecular dynamics simulations. *Macromolecules* **1994**, *27* (26), 7781-7784.
11. Chantawansri, T. L.; Yeh, I.-C.; Hsieh, A. J., Investigating the glass transition temperature at the atom-level in select model polyamides: A molecular dynamics study. *Polymer* **2015**, *81*, 50-61.
12. Patrone, P. N.; Dienstfrey, A.; Browning, A. R.; Tucker, S.; Christensen, S., Uncertainty quantification in molecular dynamics studies of the glass transition temperature. *Polymer* **2016**, *87*, 246-259.
13. Minelli, M.; De Angelis, M. G.; Hofmann, D., A novel multiscale method for the prediction of the volumetric and gas solubility behavior of high-Tg polyimides. *Fluid Phase Equilibria* **2012**, *333*, 87-96.
14. Lyulin, S. V.; Gurtovenko, A. A.; Larin, S. V.; Nazarychev, V. M.; Lyulin, A. V., Microsecond atomic-scale molecular dynamics simulations of polyimides. *Macromolecules* **2013**, *46* (15), 6357-6363.
15. Falkovich, S. G.; Lyulin, S. V.; Nazarychev, V. M.; Larin, S. V.; Gurtovenko, A. A.; Lukasheva, N. V.; Lyulin, A. V., Influence of the electrostatic interactions on thermophysical properties of polyimides: Molecular-dynamics simulations. *Journal of Polymer Science Part B: Polymer Physics* **2014**, *52* (9), 640-646.
16. Yang, Q.; Chen, X.; He, Z.; Lan, F.; Liu, H., The glass transition temperature measurements of polyethylene: determined by using molecular dynamic method. *RSC Advances* **2016**, *6* (15), 12053-12060.
17. Cousin, T.; Galy, J.; Dupuy, J., Molecular modelling of polyphthalamides thermal properties: Comparison between modelling and experimental results. *Polymer* **2012**, *53* (15), 3203-3210.
18. Younker, J. M.; Poladi, R. H.; Bendler, H. V.; Sunkara, H. B., Computational screening of renewably sourced polyalkylene glycol plasticizers for nylon polyamides. *Polymers for Advanced Technologies* **2016**, *27* (2), 273-280.

19. Li, C.; Medvedev, G. A.; Lee, E.-W.; Kim, J.; Caruthers, J. M.; Strachan, A., Molecular dynamics simulations and experimental studies of the thermomechanical response of an epoxy thermoset polymer. *Polymer* **2012**, *53* (19), 4222-4230.
20. Li, C.; Strachan, A., Effect of thickness on the thermo-mechanical response of free-standing thermoset nanofilms from molecular dynamics. *Macromolecules* **2011**, *44* (23), 9448-9454.
21. Fermeglia, M.; Pricl, S., Multiscale modeling for polymer systems of industrial interest. *Progress in Organic Coatings* **2007**, *58* (2-3), 187-199.
22. Santiago García, J. L.; Bastarrachea, M. I. L.; de Jesús Aguilar Vega, M. In *Aromatic polyamides density from molecular dynamics simulation*, Macromolecular Symposia, Wiley Online Library: 2013; pp 120-124.
23. Li, M.; Liu, X.; Qin, J.; Gu, Y., Molecular dynamics simulation on glass transition temperature of isomeric polyimide. *Express Polym Lett* **2009**, *3*, 665-675.
24. Wunderlich, B., Glass transition as a key to identifying solid phases. *Journal of Applied Polymer Science* **2007**, *105* (1), 49-59.
25. Soldera, A.; Metatla, N., Glass transition of polymers: Atomistic simulation versus experiments. *Physical Review E* **2006**, *74* (6), 061803.
26. Li, C.; Strachan, A., Molecular scale simulations on thermoset polymers: A review. *Journal of Polymer Science Part B: Polymer Physics* **2015**, *53* (2), 103-122.
27. Hossain, D.; Tschopp, M. A.; Ward, D.; Bouvard, J.-L.; Wang, P.; Horstemeyer, M. F., Molecular dynamics simulations of deformation mechanisms of amorphous polyethylene. *Polymer* **2010**, *51* (25), 6071-6083.
28. Wu, C., Simulated glass transition of poly (ethylene oxide) bulk and film: a comparative study. *The Journal of Physical Chemistry B* **2011**, *115* (38), 11044-11052.
29. Afzal, M. A. F.; Sanders, J. M.; Goldberg, A.; Browning, A. R.; Halls, M. D., Using Molecular Simulation with High-Temperature Composites Resins. In *SAMPE*, Charlotte, NC, 2019.
30. Aylwin, P. A.; Boyd, R. H., Aliphatic polyesters as models for relaxation processes in crystalline polymers: 1. Characterization. *Polymer* **1984**, *25* (3), 323-329.
31. Boyd, R. H., Relaxation processes in crystalline polymers: experimental behaviour — a review. *Polymer* **1985**, *26* (3), 323-347.
32. Gibbs, J. H.; DiMarzio, E. A., Nature of the glass transition and the glassy state. *The Journal of Chemical Physics* **1958**, *28* (3), 373-383.
33. Hoffman, J. D., Anelastic and dielectric effects in polymeric solids, N. G. McCrum, B. E. Read, and G. Williams, Wiley, New York, 1967. pp. 617. \$25.00. *Journal of Applied Polymer Science* **1969**, *13* (2), 397-397.
34. Stillingner, F. H.; Debenedetti, P. G., Glass Transition Thermodynamics and Kinetics. *Annual Review of Condensed Matter Physics* **2013**, *4* (1), 263-285.
35. Sjogren, L., Diffusion of impurities in a dense fluid near the glass transition. *Phys Rev A Gen Phys* **1986**, *33* (2), 1254-1260.
36. Yin, H.; Napolitano, S.; Schönhals, A., Molecular Mobility and Glass Transition of Thin Films of Poly(bisphenol A carbonate). *Macromolecules* **2012**, *45* (3), 1652-1662.
37. Boyer, R. F.; Meier, D. J., *Molecular Basis of Transitions and Relaxations*. Gordon and Breach Science Publishers: 1978.
38. Tesoro, G., Textbook of polymer science, 3rd ed., Fred W. Billmeyer, Jr., Wiley-Interscience, New York, 1984, 578 pp. No price given. *Journal of Polymer Science: Polymer Letters Edition* **1984**, *22* (12), 674-674.
39. Murray, G., *Handbook of Materials Selection for Engineering Applications*. Taylor & Francis: 1997.
40. Wunderlich, B., Reversible crystallization and the rigid–amorphous phase in semicrystalline macromolecules. *Progress in Polymer Science* **2003**, *28* (3), 383-450.

41. Starr, F. W.; Douglas, J. F., Modifying Fragility and Collective Motion in Polymer Melts with Nanoparticles. *Physical Review Letters* **2011**, *106* (11), 115702.
42. Soldera, A.; Grohens, Y., Local Dynamics of Stereoregular PMMAs Using Molecular Simulation. *Macromolecules* **2002**, *35* (3), 722-726.
43. Hung, J.-H.; Patra, T. K.; Simmons, D. S., Forecasting the experimental glass transition from short time relaxation data. *Journal of Non-Crystalline Solids* **2020**, *544*, 120205.
44. Ozawa, M.; Scalliet, C.; Ninarello, A.; Berthier, L., Does the Adam-Gibbs relation hold in simulated supercooled liquids? *The Journal of Chemical Physics* **2019**, *151* (8), 084504.
45. *Materials Science Suite*, Schrödinger, LLC: New York, NY, 2018.
46. *Schrödinger Release 2016-4: Desmond Molecular Dynamics System*, Schrödinger, New York, NY, 2016.
47. Lin, Y. H., Entanglement and the molecular weight dependence of polymer glass transition temperature. *Macromolecules* **1990**, *23* (25), 5292-5294.
48. Shaw, D. E.; Grossman, J. P.; Bank, J. A.; Batson, B.; Butts, J. A.; Chao, J. C.; Deneroff, M. M.; Dror, R. O.; Even, A.; Fenton, C. H.; Forte, A.; Gagliardo, J.; Gill, G.; Greskamp, B.; Ho, C. R.; Ierardi, D. J.; Iserovich, L.; Kuskin, J. S.; Larson, R. H.; Layman, T.; Lee, L.; Lerer, A. K.; Li, C.; Killebrew, D.; Mackenzie, K. M.; Mok, S. Y.; Moraes, M. A.; Mueller, R.; Nociolo, L. J.; Peticolos, J. L.; Quan, T.; Ramot, D.; Salmon, J. K.; Scarpazza, D. P.; Schafer, U. B.; Siddique, N.; Snyder, C. W.; Spengler, J.; Tang, P. T. P.; Theobald, M.; Toma, H.; Towles, B.; Vitale, B.; Wang, S. C.; Young, C. In *Anton 2: Raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer*, SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 16-21 Nov. 2014; 2014; pp 41-53.
49. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A., OPLS3: A force field providing broad coverage of drug-like small molecules and proteins. *Journal of Chemical Theory and Computation* **2016**, *12* (1), 281-296.
50. Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L.; Abel, R.; Friesner, R. A.; Harder, E. D., OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *Journal of Chemical Theory and Computation* **2019**, *15* (3), 1863-1874.
51. Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W., Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *Journal of Chemical Theory and Computation* **2010**, *6* (5), 1509-1519.
52. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* **1996**, *118* (45), 11225-11236.
53. Jorgensen, W. L.; Tirado-Rives, J., The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society* **1988**, *110* (6), 1657-1666.
54. Nosé, S., A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **1984**, *81* (1), 511-519.
55. Hoover, W. G., Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **1985**, *31* (3), 1695-1697.
56. Martyna, G. J.; Tobias, D. J.; Klein, M. L., Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **1994**, *101* (5), 4177-4189.
57. Weininger, D., SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Computer Sciences* **1988**, *28* (1), 31-36.

58. Afzal, M. A. F.; Sonpal, A.; Haghighatlari, M.; Schultz, A. J.; Hachmann, J., A deep neural network model for packing density predictions and its application in the study of 1.5 million organic molecules. *Chemical Science* **2019**, *10* (36), 8374-8383.
59. Doerr, S.; Harvey, M.; Noé, F.; De Fabritiis, G., HTMD: high-throughput molecular dynamics for molecular discovery. *Journal of chemical theory and computation* **2016**, *12* (4), 1845-1852.
60. AFZAL, M. A. From virtual high-throughput screening and machine learning to the discovery and rational design of polymers for optical applications. State University of New York at Buffalo, 2018.
61. Mohammad Atif Faiz, A.; Jarod M., Y.; George, R., *The Effect of Tacticity and Side Chain Structure on the Coil Dimensions of Polyolefins*. 2018.
62. Elton, D. C.; Boukouvalas, Z.; Fuge, M. D.; Chung, P. W., Deep learning for molecular design-a review of the state of the art. *Molecular Systems Design & Engineering* **2019**.
63. Hung, J.-H.; Patra, T. K.; Meenakshisundaram, V.; Mangalara, J. H.; Simmons, D. S., Universal localization transition accompanying glass formation: insights from efficient molecular dynamics simulations of diverse supercooled liquids. *Soft Matter* **2019**, *15* (6), 1223-1242.
64. Xie, R.; Weisen, A. R.; Lee, Y.; Aplan, M. A.; Fenton, A. M.; Masucci, A. E.; Kempe, F.; Sommer, M.; Pester, C. W.; Colby, R. H.; Gomez, E. D., Glass transition temperature from the chemical structure of conjugated polymers. *Nature Communications* **2020**, *11* (1), 893.
65. Cassar, D. R.; de Carvalho, A. C.; Zanotto, E. D., Predicting glass transition temperatures using neural networks. *Acta Materialia* **2018**, *159*, 249-256.
66. Sukumar, N.; Krein, M.; Luo, Q.; Breneman, C., MQSPR modeling in materials informatics: a way to shorten design cycles? *Journal of Materials Science* **2012**, *47* (21), 7703-7715.
67. Jha, A.; Chandrasekaran, A.; Kim, C.; Ramprasad, R., Impact of dataset uncertainties on machine learning model predictions: the example of polymer glass transition temperatures. *Modelling and Simulation in Materials Science and Engineering* **2019**, *27* (2), 024002.
68. Andjelić, S.; Scogna, R. C., Polymer crystallization rate challenges: The art of chemistry and processing. *Journal of Applied Polymer Science* **2015**, *132* (38), 42066.
69. Kovacs, A. J., La contraction isotherme du volume des polymères amorphes. *Journal of Polymer Science* **1958**, *30* (121), 131-147.
70. Dudowicz, J.; Douglas, J.; Freed, K., The meaning of the "universal" WLF parameters of glass-forming polymer liquids. *The Journal of chemical physics* **2015**, *142*, 014905.
71. Morgan, R. J.; Kong, F.-M.; Walkup, C. M., Structure-property relations of polyethertriamine-cured bisphenol-A-diglycidyl ether epoxies. *Polymer* **1984**, *25* (3), 375-386.
72. Zheng, Y.; McKenna, G. B., Structural Recovery in a Model Epoxy: Comparison of Responses after Temperature and Relative Humidity Jumps. *Macromolecules* **2003**, *36* (7), 2387-2396.
73. Montserrat, S.; Calventus, Y.; Hutchinson, J. M., Effect of cooling rate and frequency on the calorimetric measurement of the glass transition. *Polymer* **2005**, *46* (26), 12181-12189.
74. Hallensleben, M. L., Polyvinyl Compounds, Others. In *Ullmann's Encyclopedia of Industrial Chemistry*, 2000; pp 1-23.