doi.org/10.26434/chemrxiv.12250229.v1

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

Mohammad Atif Faiz Afzal, Andrea Browning, Alexander Goldberg, Mathew D. Halls, Jacob L. Gavartin, Tsuguo Morisato, Thomas F. Hughes, David J. Giesen, Joseph E. Goose

Submitted date: 06/05/2020 • Posted date: 08/05/2020

Citation information: Afzal, Mohammad Atif Faiz; Browning, Andrea; Goldberg, Alexander; Halls, Mathew D.; Gavartin, Jacob L.; Morisato, Tsuguo; et al. (2020): High-Throughput Molecular Dynamics Simulations and Validation of Thermophysical Properties of Polymers. ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.12250229.v1

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 of large molecular systems. This paper outlines an efficient MD cooling 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 experimental values were reported by Bicerano.[1] We observed good agreement of predicted T_g values with experimental observation across a wide range of polymers, which confirms the clear utility of the described workflow. 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.

File list (3)

Tg_paper_Megarun.pdf (1.57 MiB)	view on ChemRxiv • download file
Tg_paper_Megarun_SI.docx (19.28 KiB)	view on ChemRxiv • download file
Tg_paper_Megarun_SI_data.xlsx (68.67 KiB)	view on ChemRxiv • download file

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

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

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 of large molecular systems. This paper outlines an efficient MD cooling 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 experimental values were reported by Bicerano.[1] We observed good agreement of predicted T_g values with experimental observation across a wide range of polymers, which confirms the clear utility of the described workflow. 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 important 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] and Differential Scanning Calorimetry (DSC)[7]. T_g measured by different methods may also display substantial variability.[8] Computational methods have also been developed that allow calculating T_g. Bicerano[1] 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 provide 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.[9-24] The accuracy of MD predicted T_g has been evaluated by numerous groups. It is typically undertaken for small sets of polymeric systems[25] of a specific type (e.g. polyethylenes[26] and ethylene oxides[27], polyamides[10], polyimides[22, 28]). 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. In order 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 universal MD screening approach for T_g prediction. In particular, we ask if a generic good-for-all protocol can be developed to estimate reliably the glass transition in broad classes of polymeric systems. 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. 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 second order[29-32] when properties of a material have a point of discontinuity as a function of temperature, while glass transition can be considered a first-order transition and thus recognized by a change in the slope of a linear plot of a thermophysical property as a function of temperature. Differences of opinions exist regarding the exact nature of the glass transition phenomenon and whether it is a true first-order transition. Some authors[33] 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. We will examine the visible discontinuities based on a detailed temperature dependence analysis.

Unlike other materials, amorphous and semi-crystalline polymers are unique in that they undergo glass transitions, i.e. a phase transition that changes materials from glassy hard, stiff and brittle state to viscous rubbery state.[34, 35] When they do, their properties such as the coefficient of thermal expansion, the heat capacity, viscosity, refractive index, and others dramatically change. [35, 36] The onset of change can be observed and quantified using the dependence of density or specific volume on the temperature. When a phase transition occurs, all the properties mentioned above experience a shift in the temperature derivatives.[35] 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 should have both crystallization and glass transitions. Its crystalline part undergoes melting at the melting temperature, while the amorphous part experiences glass transition at T_g. However, T_g is also affected by the crystalline part but the way it is affected is not well known. Although the main 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 wholly crystalline polymers do not have a glass transition phase change detectable with standard techniques. In crystalline polymers, melting occurs. Some of the polymers in our study develop ordering and this process can be interpreted as partial crystallization. When a polymer crystallizes, it does not show a clear glass transition.

Empirical detection of glass transitions is based on two main approaches.[35, 36] 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 Bicerano[1] taken from different sources where different experimental methods were used, so scatter across the polymer systems in T_g value related to experimental technique may be present. A more detailed validation with analysis of the agreement with different experimental techniques is out of the scope of this study.

T_g calculations using molecular-level simulations have been performed since the 1990s.[9] These methods rely on the accurate prediction of the volume versus temperature behavior of polymers in both high temperature and low temperature regions. 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 easily handle T_g's that vary over 100's of degrees.

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 Bicerano[1] 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. [37, 38]

b. Simulation protocol

We built polymer chains such that the total number of atoms in a single chain is approximately 2000 atoms. 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. Subsequently, we constructed simulation cells with six polymer chains resulting in a box with approximately 12000 atoms. 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.[38, 39] The simulations use the OPLS3e force field for all the systems in this work.[36, 40-44]

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. For the polymer systems, whose expected T_g is greater than 600 K, we selected the starting temperature to be 1100 K and end temperature as 200 K. For all styrene containing polymers, the starting temperature was 1100 K and the step size was 10 K instead of 20 K. 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[45, 46] and the Martina-Tobias-Klein barostat.[47] We considered the density to be converged if the standard deviation did not exceed 5%.

To calculate T_g from MD simulations, researchers have typically used the specific volume as a function of temperature (v(T)).[9, 10, 12-25] 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 could fail. 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 limits the use of such a method in automated screening studies. Another approach to computing the T_g is by performing a hyperbolic fit to the density vs temperature data. The hyperbola fit method was developed 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. Hyperbolic fit was shown to overcome the ambiguity of the bi-linear fits and allows the automated estimation of T_g values.[11] 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 perform the T_g calculation protocol on each of these replicates and compute the T_g value by performing the hyperbolic fit. We compute the aggregate T_g and standard deviation of the polymer by taking the weighted average of T_g values from all the ten replicates, where weights are the inverse of the within simulation uncertainty as described in Patrone.[11]

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 replicates 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 1. We first converted the polymer names from the Bicerano handbook to machine readable SMILES (Simplified Molecular Input Line Entry System) strings.[48] 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 replicates 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 replicates and performing the uncertainty quantification, and 7) resubmission of T_g jobs, with a larger temperature window, if the standard deviation between replicates is more than 50 K. Additionally, a guess value for the T_g can also be provided in the process to select a narrow temperature window in the hyperbola fit. Due to the high computational costs associated with polymer MD simulations, there have been very few high-throughput screening studies in the past.[49-52] However, with the advancement in hardware and efficient implementation of GPU accelerated MD code, we are now able to perform high-throughput MD screening. Using the automated T_g workflow in this work, we performed more than 150,000 simulations (including replicates and individual temperature runs), which comprised of about 1 million ns of MD time.

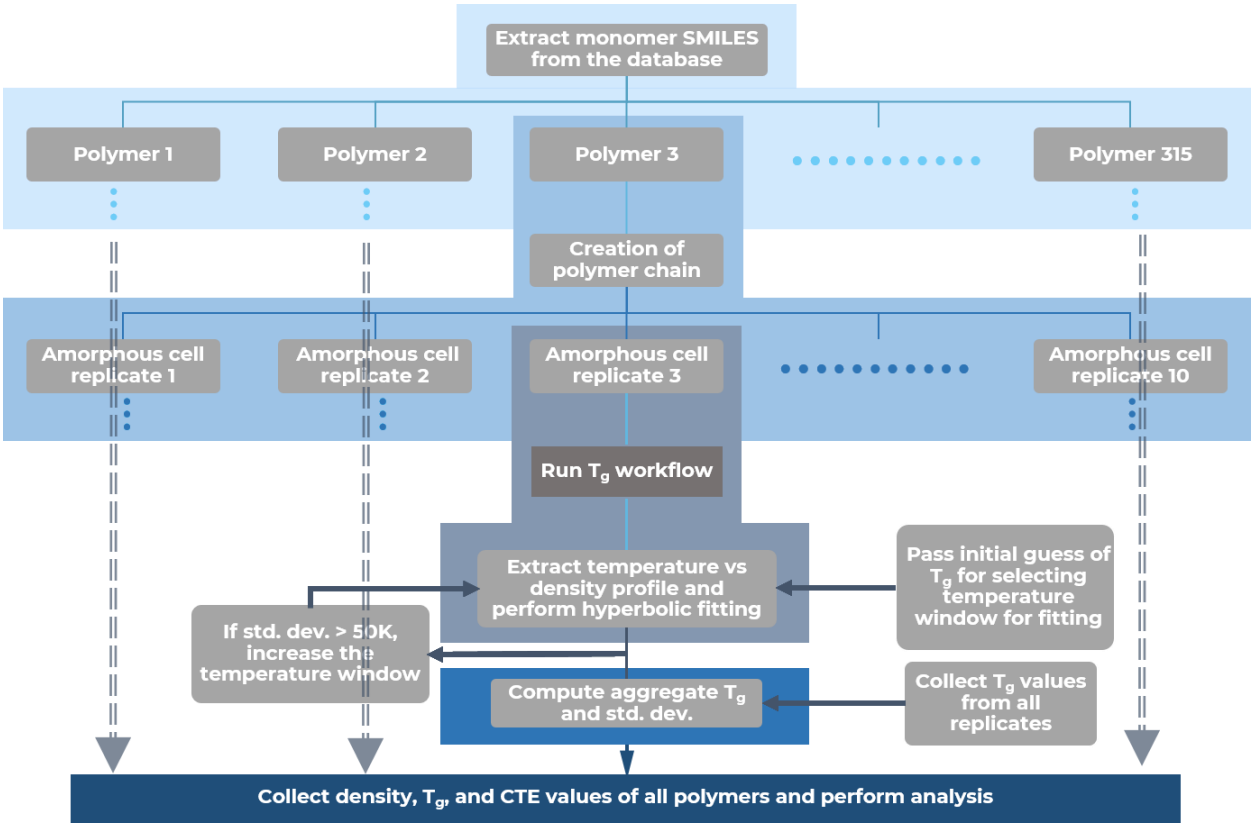

Figure 1: Automated workflow for the high-throughput screening of 315 polymers to evaluate the T_g values.

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 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.[53] Details of the simulations cell and T_g protocol for these systems are provided in the SI (Supplementary Information). It should be noted that we only use a single replicate for these simulation cells as the main 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 the polymers where we observe ordering.

f. *Coefficient of thermal expansion*

As the process of T_g computations involve 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 taking ten points from the lower temperature range and ten points from the high temperature region, respectively. We selected these ten points such that it is sufficient to provide a consistent linear region. 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 2 shows the comparison of calculated and experimental density values, where the histograms on the top x-axis and right y-axis illustrate the distribution of experimental density and calculated density values, respectively. 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 (+2.02%), 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. Some of the outliers with the prediction error of more than 10% include polychlorotrifluoroethylene, poly(vinyl methyl sulfide), poly(N-vinyl pyrrolidone), and poly(methyl a-cyanoacrylate). Further investigation is needed to understand the source of this discrepancy.

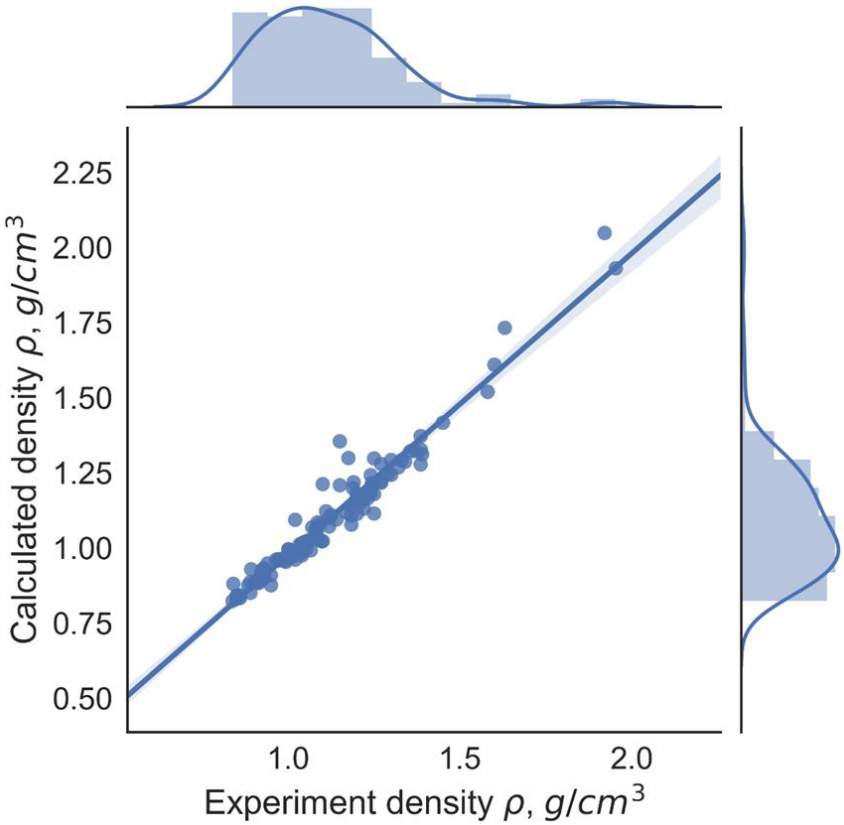

Figure 2. Predicted density values against experimental density reported in Bicerano[1] table 3.5. The blue line is the linear regression and the shaded area is the uncertainty in the linear fitting. The histograms on the top x-axis and right y-axis show the distribution of experimental density and calculated density values, respectively.

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.[11] An example of the automated hyperbola fit is shown in Figure 3 for polyoxyoctamethylene polymer. Based on the fitting, the T_g of this polymer is calculated as 238 K, which is overestimated compared to 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,[9, 54] 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,[39] 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 behavior and universal parameters for polymers.[19] 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.

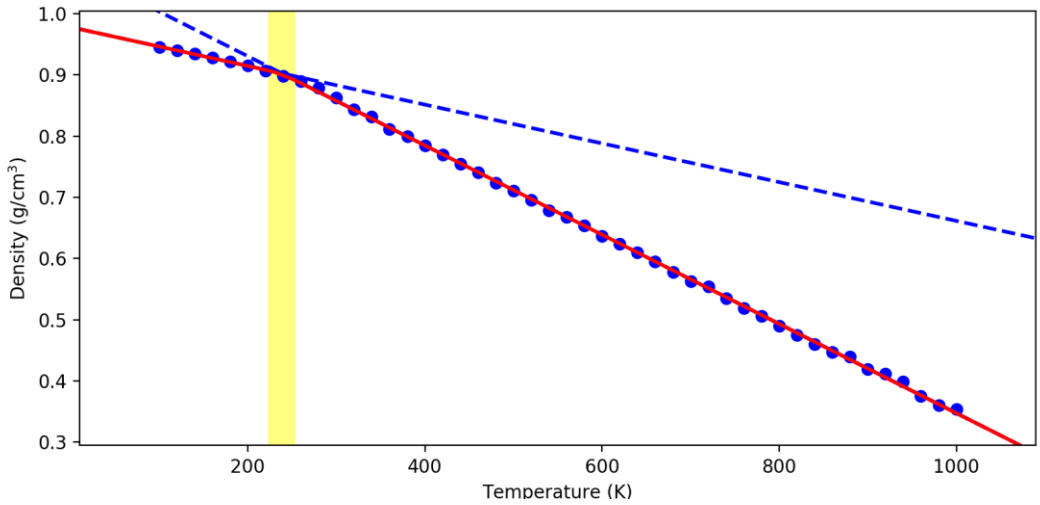

Figure 3. Hyperbola fit of density as a function of temperature to evaluate a glass transition temperature (T_g) of polyoxyoctamethylene. The red curve represents the hyperbolic curve and blue dashed lines represent the asymptotes. The yellow region represents the portion of simulated density values 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 4. We included the absolute values in the supplementary information. As expected, for most of the polymers, observed T_g values are overestimated compared to the experimental values. However, the R^2 value of 0.92 shows that the trends in the results are in good agreement 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 (calc.) + 21.08 K. We show the calibrated T_g values of all the 315 polymers in Figure 5. The prediction errors for the calibrated T_g are reasonable, with mean absolute error (MAE) of 27.5 K (7.98%) and a root mean squared error (RMSE) of 36.18 (10.92%).

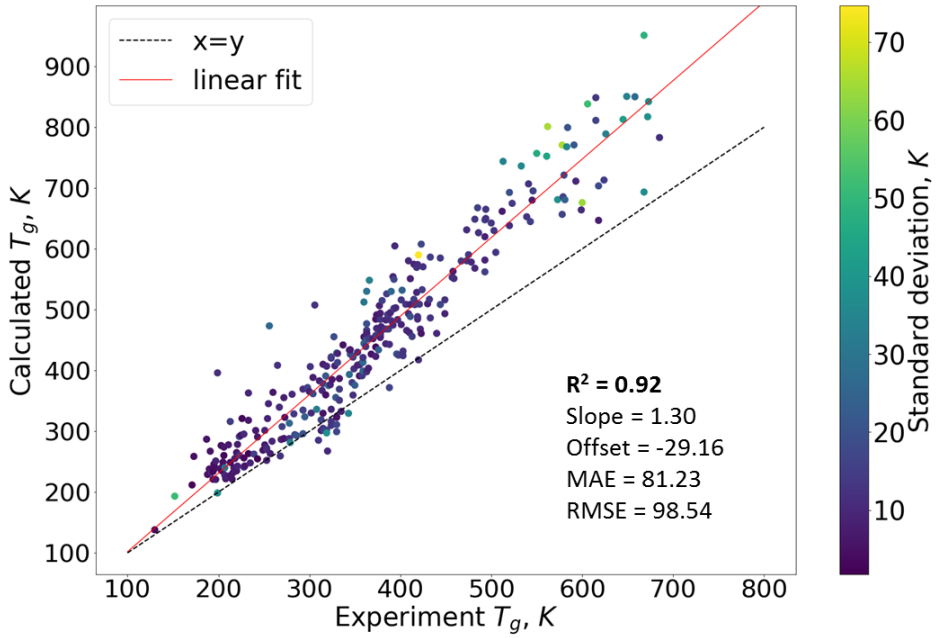

Figure 4. Predicted T_g values against experimental ones reported in Bicerano[1] Table 6.2. The color bar represents the standard deviation of prediction based on the uncertainty quantification.

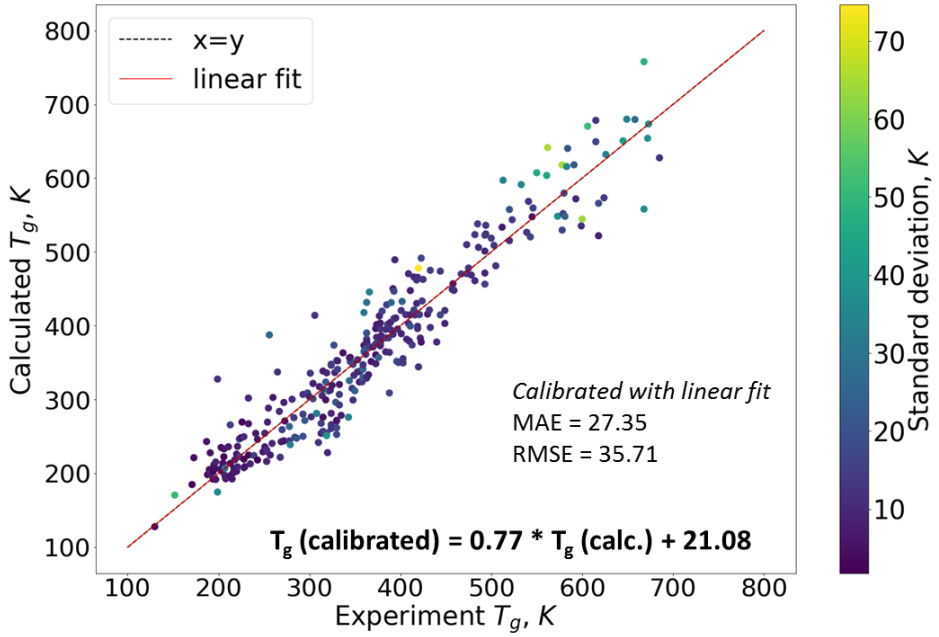

Figure 5. Calibrated T_g values against experimental ones reported in Bicerano[1] Table 6.2. The color bar represents the standard deviation of prediction based on the uncertainty quantification.

The color of the points in Figure 4 and Figure 5 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

region 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(predicted) – T_g(experiment) (Figure 6). Figure 6a shows the deviations of the calculated T_g from the experimental T_g, whereas Figure 6b shows the deviation 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.

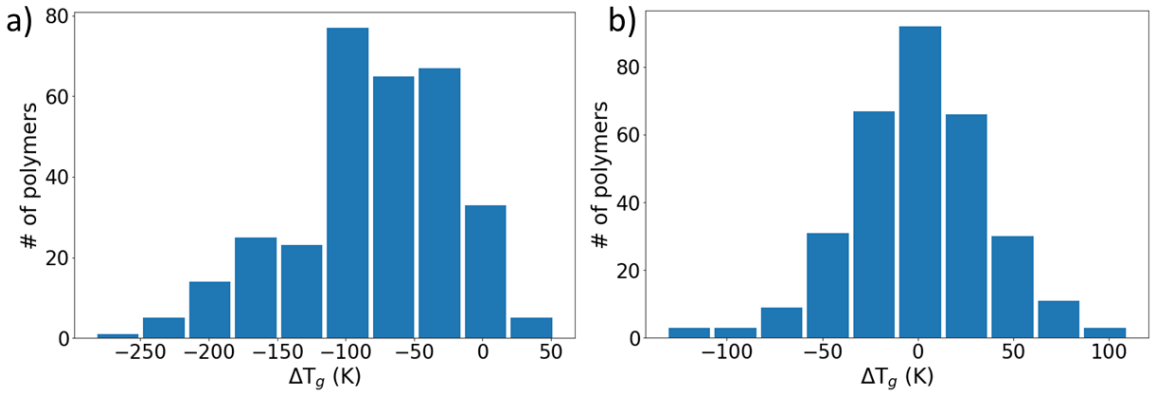

Figure 6. Histogram showing the deviations of the T_g from; a) the parity line, and b) the linear regression fit.

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,[36] the melting temperature is in the range of 380-420 K. Our calculation underestimates the onset of melting that is apparently due to the different amorphous fraction in the experimentally prepared polyethylene.

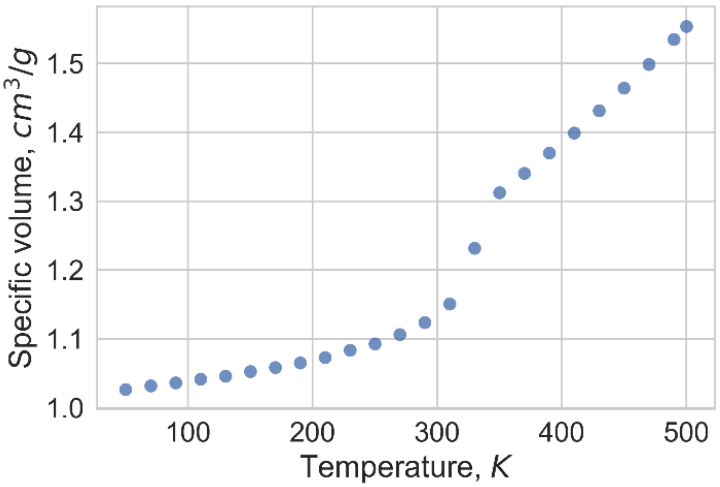

Figure 7. Specific volume as a function of temperature for linear polyethylene polymer, and b) branched polyethylene polymer. The second-order transition can be associated with the melting temperature (T_m).

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.[55] According to MD calculations, at the temperature of 600 K, the polymer chains in the unit cell are shaping 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,[29, 30] the T_g of the semi-crystalline polymer should correspond to the T_g of its amorphous part. However, the effect of the crystalline part is not obvious. According to our fitting protocol, the T_g in such a case cannot be determined unambiguously (Figure 8b).

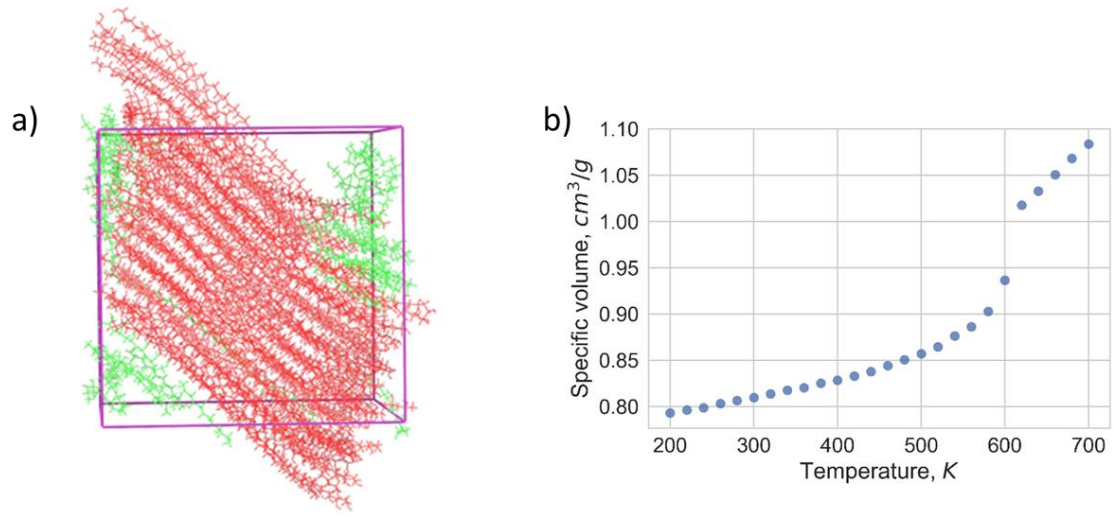

Figure 8. Snapshot of the polyvinyl alcohol polymer at 600 K (a) and specific volume as a function of temperature (b). The onset of crystallization can be identified in the plot.

Traditionally, the increase in the intermolecular association such as hydrogen bonding favors the formation of semi-crystalline phases in polymer materials.[30] Polymers that contain aromatic rings are more rigid and through stacking interactions, they are more likely to crystallize. This stacking interaction was confirmed experimentally for poly(p-phenylene terephthalamide). Chantawansri *et al*.[10] argued that the time and length scale accessible to all-atom molecular dynamics cannot capture such a phenomenon as phase separation and the formation of the crystalline regions. However, recent developments in the state-of-the-art MD simulation, especially the capability to use GPU, expands the simulation time to about 100 ns and beyond.[39] Using such simulations, we can observe the formation of the crystalline phase (Figure 9a). The onset of the ordered phase can be identified in the specific volume and temperature plot (Figure 9b), at approximately 1100 K. It should be noted that the glass transition occurrence cannot be established in such plot.

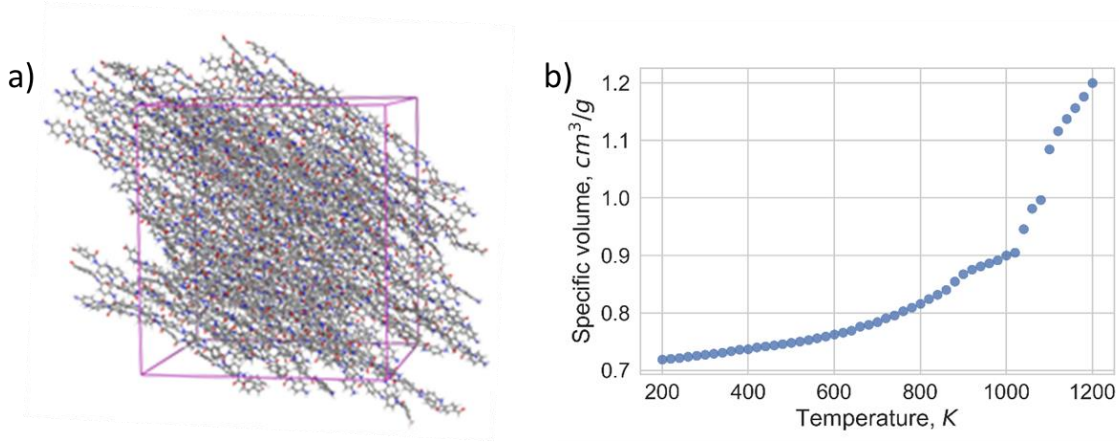

Figure 9. Image of poly(p-phenylene terephalamide) polymer at 800 K (a) and specific volume as a function of temperature (b). The onset of crystallization can be identified in the plot.

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. In Figure 10, which shows the specific volume variation for the branched polyethylene, the melting temperature is not unambiguously observed. As we increase the branching, the melting transition becomes less and 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.

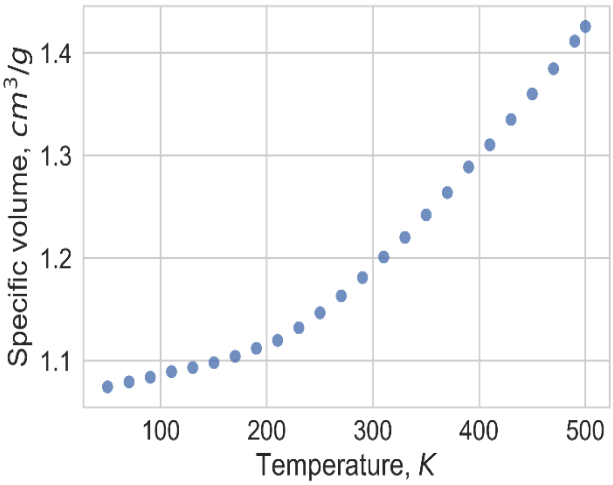

Figure 10. Specific volume as a function of temperature for branched polyethylene polymer

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 11a and 11b, respectively. Glassy CTE for these polymers ranges from $100 \times 10^{-6}\ K^{-1}$ to $500 \times 10^{-6}\ K^{-1}$, with a median value of $200 \times 10^{-6}\ K^{-1}$. We provide the calculated CTE values of the polymers in the SI. Figure 11c 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 10 replicates. Figure 11d 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. 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.

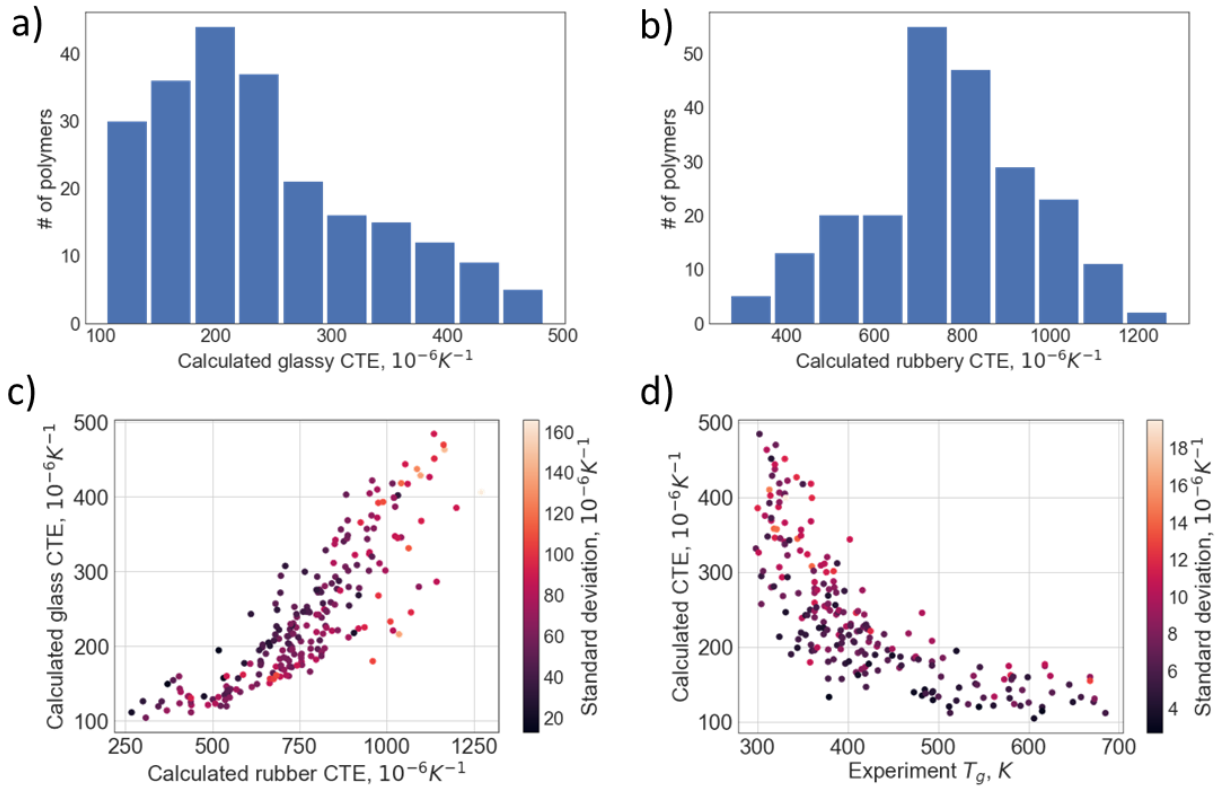

Figure 10: a) Distribution of glassy CTE values of the polymers; b) Distribution of rubbery CTE values of the polymers; c) Comparison of glassy CTE and rubber CTE values; d) Variation of glassy CTE values with the experimental T_g.

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. The validation studies show good agreement with the experimental values indicating that our model is effective and economical in evaluating the T_g of polymers. Additionally, we identified areas for improvement in the prediction of high temperature polymers and in polymers exhibiting multiple transition points.

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 point. We confirmed the partial crystallization of those polymers by looking at the geometry of those polymers below the melting temperature. We demonstrated the distinction between crystallized and non-crystalized 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 good agreement with the experiment using our T_g protocol. 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. 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. SUPPLEMENTARY INFORMATION

Electronic supplementary 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 of 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. 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.
9. Han, J.; Gee, R. H.; Boyd, R. H., Glass transition temperatures of polymers from molecular dynamics simulations. *Macromolecules* **1994,** *27* (26), 7781-7784.
10. 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.
11. 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.
12. 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.
13. 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.
14. 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.
15. 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.

16. Cousin, T.; Galy, J.; Dupuy, J., Molecular modelling of polyphthalamides thermal properties: Comparison between modelling and experimental results. *Polymer* **2012,** *53* (15), 3203-3210.
17. 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.
18. 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.
19. 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.
20. Fermeglia, M.; Pricl, S., Multiscale modeling for polymer systems of industrial interest. *Progress in Organic Coatings* **2007,** *58* (2-3), 187-199.
21. 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.
22. 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.
23. Wunderlich, B., Glass transition as a key to identifying solid phases. *Journal of Applied Polymer Science* **2007,** *105* (1), 49-59.
24. Soldera, A.; Metatla, N., Glass transition of polymers: Atomistic simulation versus experiments. *Physical Review E* **2006,** *74* (6), 061803.
25. 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.
26. 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.
27. 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.
28. 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.
29. Aylwin, P. A.; Boyd, R. H., Aliphatic polyesters as models for relaxation processes in crystalline polymers: 1. Characterization. *Polymer* **1984,** *25* (3), 323-329.
30. Boyd, R. H., Relaxation processes in crystalline polymers: experimental behaviour — a review. *Polymer* **1985,** *26* (3), 323-347.
31. 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.
32. 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.
33. Sjogren, L., Diffusion of impurities in a dense fluid near the glass transition. *Phys Rev A Gen Phys* **1986,** *33* (2), 1254-1260.
34. Boyer, R. F.; Meier, D. J., *Molecular Basis of Transitions and Relaxations*. Gordon and Breach Science Publishers: 1978.
35. 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.
36. Murray, G., *Handbook of materials selection for engineering applications*. CRC Press: 1997.
37. *Materials Science Suite*, Schrödinger, LLC: New York, NY, 2018.

38. *Schrödinger Release 2016-4: Desmond Molecular Dynamics System*, Schrödinger, New York, NY, 2016.
39. 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.; Peticolas, 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.
40. 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.
41. 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.
42. 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.
43. 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.
44. 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.
45. Nosé, S., A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **1984,** *81* (1), 511-519.
46. Hoover, W. G., Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **1985,** *31* (3), 1695-1697.
47. Martyna, G. J.; Tobias, D. J.; Klein, M. L., Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **1994,** *101* (5), 4177-4189.
48. 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.
49. 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.
50. 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.
51. 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.
52. Mohammad Atif Faiz, A.; Jarod M., Y.; George, R., *The Effect of Tacticity and Side Chain Structure on the Coil Dimensions of Polyolefins*. 2018.
53. Andjelić, S.; Scogna, R. C., Polymer crystallization rate challenges: The art of chemistry and processing. *Journal of Applied Polymer Science* **2015,** *132* (38).

54. Kovacs, A. J., La contraction isotherme du volume des polymères amorphes. *Journal of Polymer Science* **1958,** *30* (121), 131-147.
55. Hallensleben, M. L., Polyvinyl Compounds, Others. In *Ullmann's Encyclopedia of Industrial Chemistry*, 2000; pp 1-23.

High-Throughput Molecular Dynamics Simulations and Validation of Thermophysical Properties of Polymers, Supplementary Information

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

1. Protocol to evaluate onset of chain ordering

For evaluating the onset of chain ordering in the polymers, we initially constructed 6 and 24 repeat units of polymers lengths. This was followed by construction of amorphous cell structure with the number of atoms in the range of 15,000 – 20,000. The initial density of the system in the amorphous cell structure was 0.5 g/cm^3. After the amorphous structure at 0.5 g/cm^3 density was prepared, we employed an equilibration protocol. All MD simulations in the workflow were based on highly efficient Desmond code.[1, 2] The equilibration procedure consisted of minimization, 0.5ns NPT dynamics with 1fs time step at temperature of 300K and pressure of 100 atm. After this short dynamics, 10ns NPT dynamics at 300K and 1atm was carried out. At this stage, the 2 fs time step was used. These stages bring most of the considered polymers to equilibrium. If the standard deviation of density was approximately 0.003-0.004 g/cm^3, the system was considered equilibrated and it can be used for further processing. In case the standard deviation was not in the specified range, another 10ns NPT dynamics was run that allows reaching the equilibrium state within the indicated standard deviation range.

This equilibrated structure was used as the initial structure to run 40ns NPT dynamics to obtain the average density of that polymer at 300K at the atmospheric pressure. The same structure was exploited to equilibrate the system at a higher temperature to evaluate the volume versus temperature behavior using the thermophysical properties protocol within MSS. The thermophysical properties protocol generated a set of MD simulations on a model system over the temperature range of twice the experimental glass transition temperature to approximately ¼

the experimental glass transition temperature. The temperature range for polyethylene, poly(vinyl alcohol), and poly(p-phenylene terephalamide) is 50-500, 200-700, and 200-1200, respectively. We selected temperature steps of 20K and with 20 ns of NPT dynamics at each temperature. The Nose-Hoover thermostat[3, 4] was employed and for the barostat the Martina-Tobias-Klein method[5] was used. The density was considered converged if standard deviation did not exceed 5%. Due to fast moving atoms at high temperatures (>700K) the time step of 1 fs was necessary for dynamics to proceed. At lower temperatures, a time step of 2 fs was utilized.

References

1. 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.; Peticolas, 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.
2. *Schrödinger Release 2016-4: Desmond Molecular Dynamics System*, Schrödinger, New York, NY, 2016.
3. Nosé, S., A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **1984,** *81* (1), 511-519.
4. Hoover, W. G., Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **1985,** *31* (3), 1695-1697.
5. Martyna, G. J.; Tobias, D. J.; Klein, M. L., Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **1994,** *101* (5), 4177-4189.

Other files

Tg_paper_Megarun_SI_data.xlsx (68.67 KiB) view on ChemRxiv • download file