ELSEVIER

PII: S0032-3861(97)00126-2

Polymer Vol. 38 No. 18, pp. 4769–4772, 1997
© 1997 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0032-3861/97/$17.00 + 0.00

The analysis of dynamic simulations of polyethylene using the methods of 'histograms and cumulants'

P. E. Klunzinger* and R. K. Eby†
Department and Institute of Polymer Science, The University of Akron, Akron, OH 44325-3909, USA
(Received 5 July 1996; revised 27 December 1996)

The methods of 'histograms and cumulants' have been applied to extract temperature dependent information from dynamic simulations of polyethylene. The use of these methods allows one to efficiently predict the coefficients of thermal expansion, and for the latter method, even to determine their temperature dependence from relatively short simulations. Using standard force fields and a run of only 470 ps, the latter method matches the coefficients of thermal expansion to 6% for the *a* unit cell dimension, 2% for *b* and 50% for the quite small, negative, coefficient in the *c*-dimension. © 1997 Elsevier Science Ltd.

(Keywords: histograms; cumulants; polyethylene; dynamics; simulation; thermal expansion; temperature; unit cells)

Introduction

One of the primary uses of molecular modelling is in the determination of the geometries of molecular systems. Force field calculations have been shown to reproduce the unit cell of many crystals. The effect of temperature on the geometries should, in principle, also be readily calculated. Unfortunately, to calculate these changes often takes considerable computational resources. In this note, we discuss methods of calculating the coefficients of thermal expansion (CTEs) for the polyethylene unit cell. The fluctuations of the periodic box will be used in the histogram and cumulant methods to illustrate the determination of CTEs for polyethylene (PE).

One of the interesting characteristics common to many extended chain polymers is the contraction with increasing temperature of the chain axis, aligned along the fibre axis in oriented fibres. This negative CTE is of technical importance in composite parts experiencing large temperature variations[1,2]. For PE, the CTEs are $-0.9 \times 10^{-5}\,K^{-1}$, $16 \times 10^{-5}\,K^{-1}$, and $6.3 \times 10^{-5}\,K^{-1}$ along the *c*, *a* and *b*-axes, respectively, over a temperature range of 100–300 K[3–5]. Calculation of the CTEs is a challenge due to the precision needed. In the present work, the computational time was kept short to increase this challenge.

Previous molecular dynamics simulations[6] have shown that the increasing contraction of the *c* axis with increasing temperature is due to 'out of plane' bending modes (perpendicular to the chain axis). This was initially demonstrated by stimulating single chains surrounded by a cell of chains fixed at positions corresponding to experimentally observed values for the temperature of interest. In the present simulations, chains are modelled in a periodic cell. By applying periodic boundary conditions: (1) all atoms are allowed to move improving the measured statistics, (2) longer range forces can more reasonably be modelled, and most importantly, (3) the dimensions of the cell can be determined as observables from the simulation (thus removing externally fixed parameters which were present in the previous simulations).

Computation

In the simulations, eight PE chains each ten ethylene units in length were placed in a periodic cell (comprising four unit cells in the *a*–*b* projection). The dimensions of the cell were allowed to change throughout the simulation by using the Anderson–Parinello–Rahman Hamiltonian[7,8]. The 'piston' mass parameter was set to be 1/10 of the system mass. A constant thermodynamic temperature was maintained using a Nose–Hoover[9,10] method as implemented in Polygraf[11]. A time step of 1 fs was used along with the Dreiding-II forcefield[12]. A 400-ps run was done to reach equilibrium at 200 K. Thereafter data were collected for a further 242 ps with data sampled every 20 fs. Runs were also done at 205 K and 225 K for times of 153 ps and 75 ps, respectively. The time average dimensions of the unit cell in these simulations are shown in *Table 1*. The uncertainties are those associated with the dimensional averages over periods of time between uncorrelated data points.

The row labelled static is of a minimized structure of the super cell. It should be noted here that the force field has been parameterized to give 'correct' geometries when minimized. Of secondary importance in the force field developments is the prediction of vibrational frequencies. As an over generalization it can be said the centres of the potential wells in the force field specify the geometry, while the force constants and functional form of these wells determine the vibrational frequencies. In this molecular dynamics study, we are examining the CTE which is more sensitive to the latter and are not concerned with the fact that the centres of the potential well are fitted to time-averaged data.

Analysis

The simplest method of calculating the CTEs is to determine the average dimensions of the unit cell at different temperatures. The results of this calculation are shown in *Table 2*. The main conclusion is that there is not enough precision to permit determination of the CTEs.

* Present address: Wavefuntion, Inc., Irvine, CA 92715, USA
† To whom correspondence should be addressed

Table 1 Calculated periodic supercell dimensions (in angstroms)

Temp.	*a* Axis	*b* Axis	*c* Axis
200 K	7.4443 ± 0.0005	4.8581 ± 0.0003	2.65414 ± 0.00001
205 K	7.48 ± 0.03	4.85924 ± 0.00005	2.65392 ± 0.00003
225 K	7.49 ± 0.02	4.8594 ± 0.0004	2.65391 ± 0.00003
Static[a]	7.3215	4.8592	2.6514
Experimental 296 K	7.217	4.901	2.546

[a] A minimized state, no motion

Table 2 Calculated CTEs from multiple simulations at different temperatures (in $10^{-5}\,K^{-1}$)

Temperature range	*a* Axis	*b* Axis	*c* Axis	Time[a] (ps)
200–205 K	96 ± 80	5 ± 1	−1.7 ± 0.2	395
200–225 K	25 ± 10	1.1 ± 0.4	−0.35 ± 0.05	317
205–225 K	7 ± 7	0.2 ± 0.4	−0.02 ± 0.08	228
Average	43 ± 32	2.1 ± 0.6	−0.68 ± 0.11	470

[a] The time shown comprises the total time for the two simulations at the temperatures given in column 1

The reason is that the fluctuations of the unit cell due to the usual temperature induced vibrations are much larger than the small changes in the time average. The uncertainties in the numbers mask changes due to increasing temperature. Since the values of the experimental CTEs do not change greatly with temperature in the range used, the averages of the calculated values of the CTE's for the different temperatures were also considered[3,4]. These too are unsatisfactory. Larger systems and simulations of sufficient length in time to sample phase space adequately are needed to bring the numbers into agreement with experimental values.

Histogram method

One weakness with the method above is that it examines only a small fraction of the data (time average values of the unit cell) collected in a dynamics simulation. Fluctuations in energy have been used to determine heat capacity of systems. It is well known that, in principle, the temperature dependence of a property can be extracted from simulations at a single temperature[13]. However, in the conventional analysis of a simulation the information contained in such fluctuations is often discarded. In the method of histograms[14] the energy distribution (which to a first approximation is a Gaussian that becomes a delta-function in the thermodynamic limit) is modelled through a population histogram $H(U_i)$ (the number of sampled states out of a total N_1 measured states that have an energy $U_i \pm \Delta U/2$) is determined. The density of states $W(U_i)$ at energy U_i can then be directly calculated. A correlation of the length [or more generically any property $A(U_i)$] dependent on the energy of the system can also be determined. Notice that both the density of states and $A(U_i)$ are independent of temperature and can be calculated from a simulation at a single temperature. The temperature dependence of $A[A(T)]$ can be calculated from these two functions via

$$A(T) = \Sigma\, A(U_i) H(U_i)\, e^{(-\delta\beta) U_i} \qquad (1)$$

Here $\delta\beta$ is the change in inverse temperature $(1/kT)$ between the simulation and temperature of interest.

An example of $H(U_i)$ calculated from the 200 K simulation and expressed as a percentage after normalization is shown in *Figure 1*. Also shown are shifted histograms, $H(U_i)\, e^{-(\delta\beta) U_i}$ for two other temperatures. There is significant noise in the wings of the curves due to poor sampling which limits the applicability of this method to a small temperature range. This noise leads to the spikes and other irregularities in the shifted curves. Nevertheless, in this small range, the nine expansion coefficients calculated (*Table 3*) are, for the most part, in

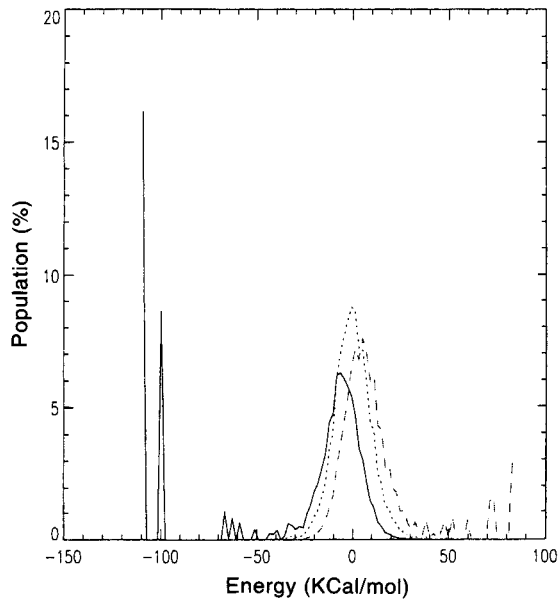

Figure 1 Three population histograms, for the temperatures 195 K (solid line), 200 K (dotted line) and 205 K (dashed line), are shown. The 200 K histogram is from the original data, the other two are calculated as discussed in the text

Table 3 Calculated CTEs from single simulations using the histogram method[a] (in $10^{-5}\,K^{-1}$)

Simulation temperature	*a* Axis	*b* Axis	*c* Axis	Time (ps)
200 K	10.1	9.3	−0.065	242
205 K	21.2	2.6	−1.5	153
225 K	14.7	7.4	−0.26	75
Average	15.3	6.4	−0.6	470

[a] Calculated with a temperature range, ΔT, of 0.2 K

better agreement with the experimental values than those based on average dimensions (*Table 2*).

The values of CTE converge much more rapidly in a given simulation making it possible to do multiple simulations from different starting configurations rather than one long run. In this context, note that the averages of the three simulations are in better general agreement with experiment than the averages given in *Table 2*.

Cumulant method

The main weakness in the histogram method can be seen in the wings of the histogram shown in *Figure 1*. In the equation above, the wings are treated as having the same accuracy as the peak of the curve even though there is much more noise in the values due to under sampling. The cumulant method has been introduced to overcome many of these obstacles when the system under study is not near a transition[15]. The central idea is to model the distribution function using a cumulant expansion instead of a histogram. (Another type of expansion also has been proposed[16].) This is justified by the fact that in a single phase region the distribution function is often nearly Gaussian and can be described by only the few lowest-order cumulants. The CTE for a side of length A can be determined using generalized cumulants[17] from

$$\Delta A = \langle A \rangle_{\beta} - \langle A \rangle_{\beta_0} = \sum_{n=0} \frac{(-\delta\beta)^{n-1}}{n!} \langle (\delta U)^n \delta A \rangle_c \quad (2)$$

Table 4 Calculated CTEs from single simulations using the second order cumulant (in $10^{-5}\,K^{-1}$)

Simulation temperature	*a* Axis	*b* Axis	*c* Axis	Time (ps)
200 K	10 ± 1	9.1 ± 1	−0.04 ± 0.4	242
205 K	21 ± 2	2.5 ± 1	−1.0 ± 0.4	153
225 K	14 ± 2	7.5 ± 2	−0.28 ± 0.4	75
Average	15	6.4	−0.44	470

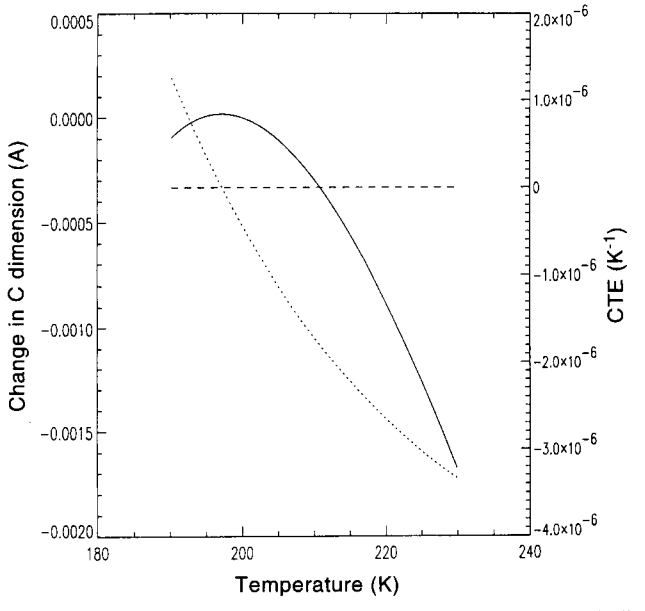

Figure 2 The change in the *c* dimension of the polyethylene periodic cell (solid line) plotted as a function temperature using three terms in the cumulant expansion of the 200 K simulation. The CTE (dotted line) is also calculated and plotted against the right axis. The dashed line indicates a value of zero for the CTE

$$CTE = \frac{1}{\langle A \rangle} \sum_{n=1} \frac{-1(-\delta\beta)^{n-1}}{T^2(n-1)!} \langle (\delta U)^n \delta A \rangle_c \quad (3)$$

where $\langle (\delta U)^n \delta A \rangle_c$ are generalized cumulants of energy and the length A. Keeping only one term in the expansion coefficient given in equation (3) is equivalent to calculating the CTE from fluctuations as presented by Allen and Tildesley[18]. Using this equation, yielded the CTE values shown in *Table 4*. Here the CTE is calculated at the simulation temperature with $\delta\beta = 0$, and therefore, only one term is used in the expansion. The uncertainty shown in the table is calculated from the second cumulant.

The length A (and the CTE) are smooth functions of $1/T$ in the cumulant approach and are thus free of many of the unphysical results that can be introduced by a discrete histogram. A plot of the predicted change in the *c*-axis as well as the CTE is shown in *Figure 2*. Using the cumulant method in the above three simulations we have been able to keep three terms for the cumulant expansion in equation (2). The first is simply the time average of $A(\langle A \rangle)$. The second is directly related to the CTE at the simulation temperature. The third term is a measure of the change in CTE as a function of temperature.

Examining the equation for CTE we see that there is a built in temperature dependence even when keeping only one term in the expansion. The sign of this dependence is, however, the opposite of that expected. In experiments with polyethylene, it has been observed that as temperature increases the CTEs of the *a* and to a lesser extent of the *b* axes of a well crystallized sample tend to become more positive while that for the *c* axis tends to become more negative[3,4]. By including a third term, $\langle (\delta U)^2 \delta A \rangle_c$ in the expansion given in equation (2) [the second term in equation (3)], the temperature dependence of CTE can be modelled better.

For example, over a temperature range of 200–225 K the experimental values for the CTE of the *a*, *b* and *c* axis changes by 1.2, 0.12 and −0.2, respectively (in units of $10^{-5}\,K^{-1}$). Using only one term in equation (2) yields predicted changes in the CTEs of −4.4, −0.5 and 0.3 from the 225 K run and −3.6, −1.4 and 0.01 from the 200 K run. As noted above, these values have the opposite sign to that seen experimentally. Including the third cumulant in equation (2) [the second term in equation (3)] yields more reasonable values of 0.49, 0.54 and −1.1 from the 225 K run and 1.20, 0.23 and −0.25 from the longer 200 K run. These have the correct signs and more nearly the correct magnitudes.

As the temperature goes to zero, the CTE is expected also to tend to zero. *Figure 2* shows the CTE (dotted line) to become zero at a much higher temperature. Keeping only three terms in equation (2) [two in equation (3)] biases the location of the zero point to be close to the simulation temperature.

Conclusion

With the use of cumulants, we are able to extract the temperature dependence of a solid state system by examining the fluctuations of the system. 'Brute force' methods of determining the CTE from multiple temperature runs are not able to determine the values as well and are not able to distinguish even the sign of the change of the CTE with temperature for the axes of PE with dynamics runs of the same time length. The use of cumulants can, on the other hand, determine not only more accurate values of

CTEs but also the temperature dependence, often with less computational effort and shorter simulations. The ability to use shorter simulations allows one to use multiple runs which may sample more of phase space than a very long, single run.

Acknowledgements

This work was sponsored in part by the Edison Polymer Innovation Corporation. Some work was carried out by R. K. Eby while at the Max Planck Institut für Polymerforschung, Mainz, Germany on an Alexander von Humboldt Senior Prize.

References

1. Liu, J. M., *Appl. Phys. Lett.*, 1986, **48**(7), 469.
2. Weinhold, P. D., Eby, R. K. and Liu, J. M., *SAMPE Quarterly*, 1993, **24**, 43.
3. Klunzinger, P. E., Ph.D. thesis, University of Akron, Polymer Science, 1993.
4. Davis, G. T., Eby, R. K. and Colson, J. P., *J. Appl. Phys.*, 1970, **41**, 4318.
5. Dadaboyev, G. and Slutsker, A. I., *Polymer Science USSR*, 1982, **24**, 1837.
6. Klunzinger, P. E., Green, K. A., Eby, R. K., Farmer, B. L., Adams, W. W. and Czornyj, G., in *Tech. Papers XXXVII.* Society of Plastics Engineers, Brookfield, CT, 1991, pp. 1532–1556.
7. Anderson, H. C., *J. Chem. Phys.*, 1980, **72**, 2384.
8. Parinello, M. and Rahman, A., *J. Appl. Phys.*, 1981, **52**, 7182.
9. Nose, S., *J. Chem. Phys.*, 1984, **81**, 511.
10. Hoover, W. H., *Phys. Rev. A*, 1985, **31**, 1695.
11. *Polygraf Reference Manual Version 3.0.* Molecular Simulation, Inc., Waltham, MA, 1992.
12. Mayo, S. L., Olafson, B. D. and Goddard III, W. A., *J. Phys. Chem.*, 1990, **94**, 8897.
13. McDonald, I. R. and Singer, K., *Discuss. Faraday Soc.*, 1967, **43**, 40.
14. Phillpot, S. R. and Rickman, J. M., *J. Chem. Phys.*, 1991, **94**, 1454.
15. Rickman, J. M. and Philpot, S. R., *Phys. Rev. Lett.*, 1991, **66**, 349.
16. Amadei, A., Apol, M. E. F., Di Nola, A. and Berendsen, H. J. C., *J. Chem. Phys.*, 1991, **94**, 1454.
17. Ryogo Kubo, *J. Phys. Soc. Japan*, 1962, **17**, 1100.
18. Allen, M. P. and Tildesley, D. J., *Computer Simulation of Liquids.* Clarendon Press, Oxford, 1987, pp. 51–54.