Received: 15 March 2023 | Revised: 20 April 2023 | Accepted: 21 April 2023

DOI: 10.1002/jcc.27128

RESEARCH ARTICLE

Journal of COMPUTATIONAL CHEMISTRY WILEY

Triacylglyceride melting point determination using coarse-grained molecular dynamics

Robert J. Cordina[1,2] | Beccy Smith[1] | Tell Tuttle[2]

[1]Mondelēz UK R&D Ltd., Birmingham, UK

[2]Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow, UK

Correspondence
Tell Tuttle, Department of Pure and Applied Chemistry, University of Strathclyde, 295 Cathedral Street, Glasgow G1 1XL, UK.
Email: tell.tuttle@strath.ac.uk

Abstract

This study is carried out using the COGITO force field to determine whether the thermodynamic melting point of pure triacylglyceride crystals can be predicted using molecular dynamics simulations. The triacylglycerides used in this study are both saturated and unsaturated, as well as symmetrical and asymmetrical, to test the robustness of both the force field and the direct heating methodology described in this paper. Given the nonequilibrium nature of a melting system, a larger number of simulations are required to ensure that the results are sufficiently converged, that is, with little fluctuation and a small confidence interval. The study also highlights the importance of the presence of defects, in this case as voids, to lower the melting nucleation energy barrier of the crystals and avoid superheating of the systems being tested. The size of these defects is much larger than what would be found in a physical crystal, however, the simple and robust procedure that was developed allows the accurate prediction of melting points of the different triacylglycerides.

KEYWORDS
COGITO, melting, phase transition

1 | INTRODUCTION

The solid-to-liquid phase transition temperature is an important macroscopic property of any material, and the simulation of this phase transition has been the research focus of a number of groups,[1–8] who have attempted to simulate the melting of various metals or molecules. Determining the melting point computationally using molecular dynamics (MD), however, is not trivial. The force field (FF) being used, and how this has been parameterized, that is, whether melting point has been part of the parameterization efforts, or not, has a significant impact on being able to predict melting points. Having a FF which is able to predict liquid and solid macroscopic properties does not imply the FF's ability to predict phase transition,[3] and hence any such simulations need to be validated thoroughly.

A common problem encountered when trying to determine the melting point of a crystal using MD is overheating (or superheating) of the system.[2,4,7,9] This is due to the system being a perfect crystal, and thus having no surface, defect or interface, or alternatively, superheating may be caused by a very high heating rate compared to empirical measurements or, the influence of the FF potential.[7] The issue of having a defect, or void, within a crystal when carrying out MD simulations has been studied extensively[1,2,4,8,9] and shown to be critical to obtaining a good determination of the melting point using MD. Having such void defects in a crystal allow for the fact that physical crystals are never perfect, and that boundaries and lattice defects reduce the possibility of crystal superheating.[2] Superheating can be mitigated by a slower heating rate, however, completing a simulation at heating rates approximating those obtainable in a lab, over a similar temperature range, is not feasible as the computation time would be too long.

Different approaches have been proposed to determine the simulated melting point. The first is a direct approach method, that is, one

where the system is heated until phase transition is observed, mimicking an experimental procedure. The second approach is to use free energy methods, with the latter being deemed to be more accurate than the former, however, it is also more complicated to carry out,[10] and can be very difficult if the solid–liquid coexistence conditions are not known in advance.[2] Given this, the direct method can be preferable as it is simpler, does not require previous knowledge of the system, and can still provide an accurate estimate of the thermodynamic melting point.[2]

Triacylglycerides (TAGs) are used in a wide variety of industries. Knowing the melting point of these compounds is very important to their processing and use. In this study, we have attempted to simulate the melting points of a range of pure TAGs with the recently developed COGITO FF,[11] using the direct heating method. The COGITO FF is a coarse-grained FF specifically parameterized to simulate the macroscopic properties of TAGs in different phases. The methodology provides a robust technique to determine this critical physical property, opening up the possibility of predicting the melting point of TAGs which are not widely available or for which the melting point is unknown.

2 | METHODS

2.1 | MD simulation settings

All simulations have been carried out using GROMACS[12] 2021.3 with the COGITO FF.[11] All equilibrations were done using an NPT ensemble, using a time-step of 25 fs and a pressure of 1.01325 bar using a v-rescale thermostat and a Berendsen barostat. Anisotropic pressure coupling, with a compressibility of 1×10^{-5} bar^{-1} in the x, y and z directions was used. Temperature coupling was set at 1 ps, while pressure coupling was set at 10 ps. The cut-off scheme was set to Verlet, with the Coulomb and vdW cut-off distances set to 1.1 nm. The vdW-modifier was set to Potential-shift and the electrostatics (coulomb type) was set to Particle-Mesh Ewald (PME) (PME order = 4). All linear heating rates were set to 0.5 K/ns (5×10^8 K/s).

2.2 | Perfect crystal generation

Pure TAG crystals were built by stacking unit cells of the TAG in all three dimensions. In the case of 1-palmitoyl-2-oleoyl-3-stearoyl-*sn*-glycerol (*sn*-POSt), 1,3-dipalmitoyl-2-oleoyl-*sn*-glycerol (*sn*-POP), 1,3-distearoyl-2-oleoyl-*sn*-glycerol (*sn*-StOSt) and 1,3-dicapric-2-oleoyl-*sn*-glycerol (*sn*-COC), where each unit cell is made up of 4 molecules, crystals were built by stacking 10, 2 and 10 unit cells in the *a*, *b* and *c* directions, respectively, to give a perfect crystal of 800 molecules (Figure 1A). In the case of 1,2,3-tristearoyl-*sn*-glycerol (*sn*-StStSt, or tristearin), where a unit cell is made up of 2 molecules, crystals were built by stacking 4, 10 and 10 unit cells in the *a*, *b* and *c* directions respectively (Figure 1B).

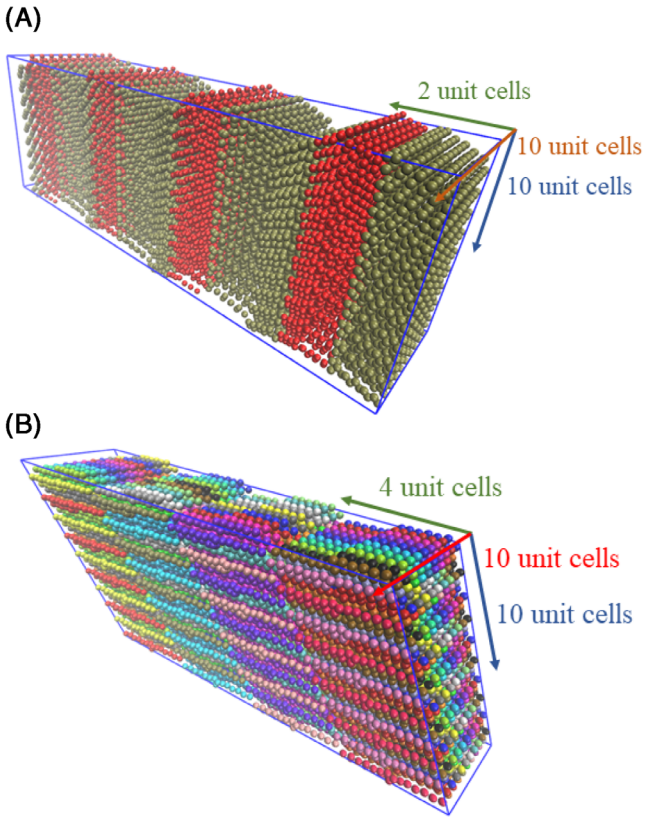

FIGURE 1 Figure showing the unit cell stacking patterns for the perfect crystals of (A) *sn*-POP (the *sn*-COC, *sn*-POSt and *sn*-StOSt crystals were built using the same unit cell stacking pattern), and (B) *sn*-StStSt.

The coarse-grained (CG) representations of *sn*-POSt, *sn*-POP, *sn*-StOSt and *sn*-StStSt were built by mapping the XRD crystalline structures[13–15] as per Cordina et al.[11] In the case of *sn*-COC, given that no crystal structure was available, a crystal was built using the CG representation of *sn*-POP as the starting point. The *sn*-COC unit cell was built by reducing the number of beads on the capric fatty acid chains, moving the molecules close to each other, using the *sn*-POP distances as a guideline, and reducing the unit cell dimensions accordingly.

2.3 | Crystals with void generation

The simulation box at any void size and type was varied between each individual simulation. These were built by starting with a perfect crystal of 800 molecules (Figure 1, and exemplified in Figure 2A), from which a molecule was chosen at random. Further molecules were then chosen at random until the required number of molecules was reached. In the case of "pit" voids the chosen molecules could be in any position within the crystal, with no relational proximity of the missing molecules except for those occurring through the random choice of molecules (Figure 2B).

In the case of "crack" voids a "seed" molecule was chosen at random. The subsequent choice of molecules to remove were chosen by

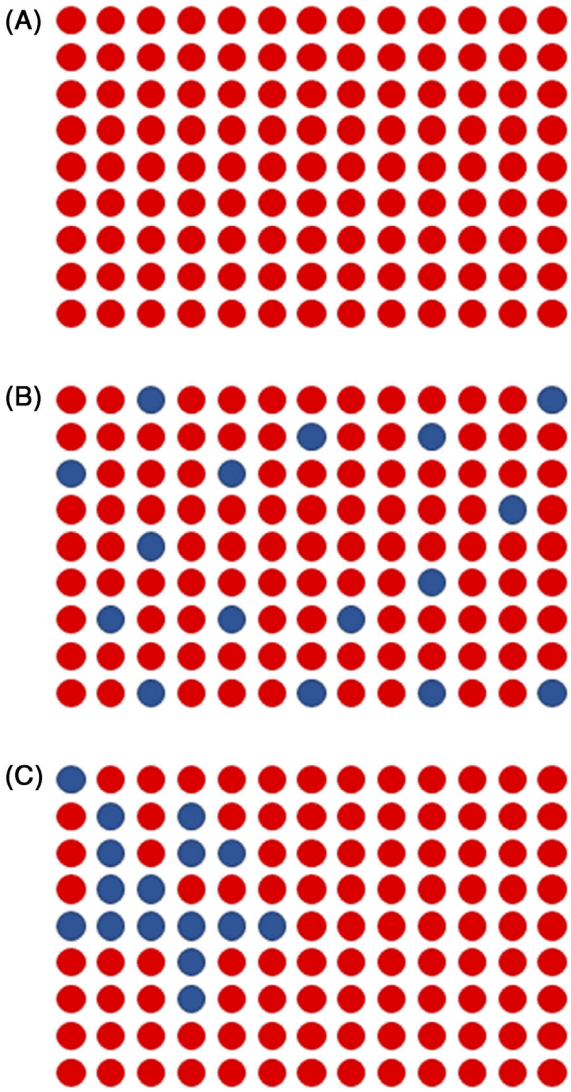

FIGURE 2 (A) 2-dimensional representation of a perfect molecular crystal; (B) crystal with 16 pit voids (blue dots = missing molecules); (C) crystal with a 16-molecule crack void.

randomly choosing a molecule which was up to one molecule away in all three dimensions from an "anchor" molecule already chosen to be removed from the crystal. The "anchor" molecule was again chosen at random each time the next molecule to be removed was to be determined, until the required number of molecules to remove were chosen. This resulted in a range of crack voids which could be anywhere in the crystal and with any shape, however, all the chosen molecules were in close relational proximity (Figure 2C). Void generation was done in an automated fashion by using a custom Python script (see Data S1).

2.4 | Melting point determination

Melting point onset for any given simulation was determined as follows:

1. the Potential energy over the trajectory was extracted using the gmx energy program in GROMACS,
2. the best-fit line (gradient and intercept) was calculated using linear regression on the data from 5% to 30% of the trajectory ("the sample set") versus temperature (°C), assuming that the temperature increase is perfectly linear with time,
3. the predicted Potential energy values over the whole trajectory/temperature range were calculated using the linear regression parameters found in the previous step,
4. the sum of the square of errors for just the sample set, that is, $\sum_{i=1}^{n}(actual_i - predicted_i)^2$, was determined,
5. the standard deviation of the sample set used for linear regression was determined, where $stdev = \sqrt{\frac{sum\ of\ square\ of\ errors}{N-2}}$,
6. the Predicted Interval as $PI = z^*{\cdot}stdev$ was determined (where $z^* = 1.96$, which statistically covers 95% of all points from the mean),
7. the actual intervals as best-fit line prediction +/− PI were calculated,
8. the melting point onset was then determined by starting from the lower temperature and iterating over the Potential energy datapoints until the number of Potential energy datapoints being higher than their respective upper PI was more than 95% of the Potential energy datapoints remaining. (all of the above was automated using a custom Python script; see Data S2)

2.5 | TAG/crystal melting temperature determination

The melting temperature for any given TAG and void size was determined as the average of the melting point onset of the total simulations at that void size. The 95% confidence interval (CI) was calculated as $CI = \frac{z^*{\cdot}\sigma}{\sqrt{N}}$ where $z^* = 1.96$ (statistical factor), σ = standard deviation of all melting point onsets for any given TAG and void size, and N = number of simulations.

The melting temperature of any given TAG was determined by varying the void size in the crystal. A plot of average melting point onset versus void size was then plotted, with the melting temperature of a TAG being determined to be that at the plateau region of the plot. A decision tree on how to choose the void size and temperature ranges is given in the Data S3.

3 | RESULTS AND DISCUSSION

3.1 | Impact of voids on melting point

To avoid superheating of the crystal, and better mimic physical crystals, a defect must be created in the crystal. A defect, be it a void, a dislocation in the crystal, or an interstitial defect, reduces the melting nucleation free energy barrier[10] which means that the nucleation rate will rise. In this study the defects were introduced as voids. Specifically, two types of voids were investigated; pits (randomly chosen

single molecules removed from the perfect crystal) and cracks (molecules chosen for their relational proximity to other already-chosen molecules) (see Computational Methods section).

The first TAG to be investigated was β_2 *sn*-POSt, where pit and crack voids were created in the perfect crystal, and heated. A detailed explanation of how the void sizes and temperature range to use for each void size was determined is provided in Data S3. Briefly, initial simulations revealed a trend consistent with what was previously described by Eike et al.[2] with the results showing a decrease in the melting point of the crystal with increasing void size (Figure 3). The decrease continues until a plateau is reached (in the case of β_2 *sn*-POSt this occurs at a 40-molecule void size), after which there is a sudden decrease in the melting point due to mechanical instability. The melting temperature of the crystal was taken to be the temperature at the plateau before the drop.[2] (see Data S4 for full melting point results).

Comparing the plots for β_2 *sn*-POSt, the melting onset for pit voids was observed to be higher than that for crack voids by approximately 10°C at any given void size (Figure 3). In this case, it is clear that the free energy barrier is higher in the cases where pit voids where used. Given these results, with the very close agreement of the simulated melting point of β_2 *sn*-POSt when using crack voids to the empirical melting point, as well as the fact that a crack void mimics defects found in crystals more closely than missing individual molecules, the remainder of the simulations, and results reported, were all carried out using crack voids.

3.2 | Accuracy of the determined melting point by bootstrap statistics

In a melting simulation on any given crystal with a void, given the nonequilibrium nature of the system, the determined melting point onset varies between repeats, even when using the same starting configuration. Given this, a number of simulations are needed to determine an average melting point. As more simulations are carried out, the changing calculated melting point (Figure 4A), and its standard error (Figure 4B), at any given void size, can be determined by using bootstrap statistics.[16,17] This statistical methodology creates numerous sub-samples by repeated resampling from a data population (in this case, using 10,000 sub-samples), thus allowing for the calculation of a population mean, standard error and confidence interval. Bootstrap statistics were chosen to minimize any disproportionate effect of an outlier on the calculation of the average melting point onset. The difference in average melting point onset using a simple arithmetic mean and bootstrap statistics turned out to be negligible.

Bootstrapping, however, was still useful in determining the minimum number of simulations required. As can be seen, the bootstrap error decreases and plateaus, while the average melting point fluctuations are reduced, with an increasing number of simulations. To determine the required number of simulations to obtain an accurate melting point at a specific void size, 250 simulations were carried out using a β_2 *sn*-POSt crystal with a 40-molecule crack void.

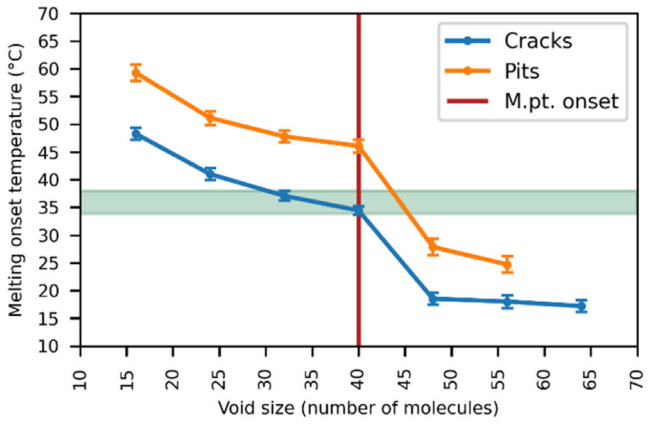

FIGURE 3 Plot of melting point of β_2 *sn*-POSt with increasing void size with pit and crack voids. Circles = determined average melting point onset; error bars = 95% CI; green shaded area = empirical melting point range.

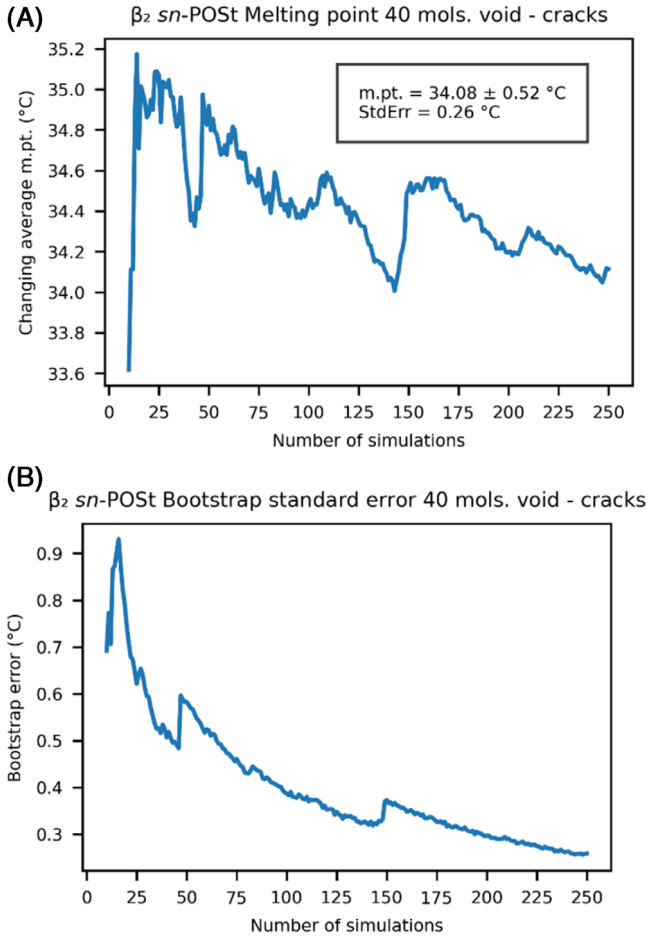

FIGURE 4 (A) Plot of melting point as determined using bootstrap statistics, with increasing number of simulations of a β_2 *sn*-POSt crystal with a 40 molecule crack void. m.pt. = average melting point ±95% confidence interval. StdErr = bootstrap standard error when using 250 datapoints. B) Plot of bootstrap standard error with increasing number of simulations.

As can be seen in Figure 4A, the calculated melting point had stabilized at around 140 simulations. A jump can then be observed, with the average melting point showing only small fluctuations after that. This jump was due to a small number of simulations with a higher melting point onset between the 140th and 150th run. Notwithstanding this, the average melting point bootstrap error increase was very small (less than 0.1°C—Figure 4B), while the change in average melting point was around 0.5°C (Figure 4A). Running more than 150 simulations did not result in an appreciable decrease in the bootstrap error or a large change in the average melting point. The average melting point for the β_2 *sn*-POSt crystal with a 40-molecule crack void calculated over 150 and 250 runs was determined to be 34.4 and 34.1°C, respectively, showing that the increase in the number of runs did not change the predicted melting point appreciably. Hence, all further melting point onsets were determined using data from

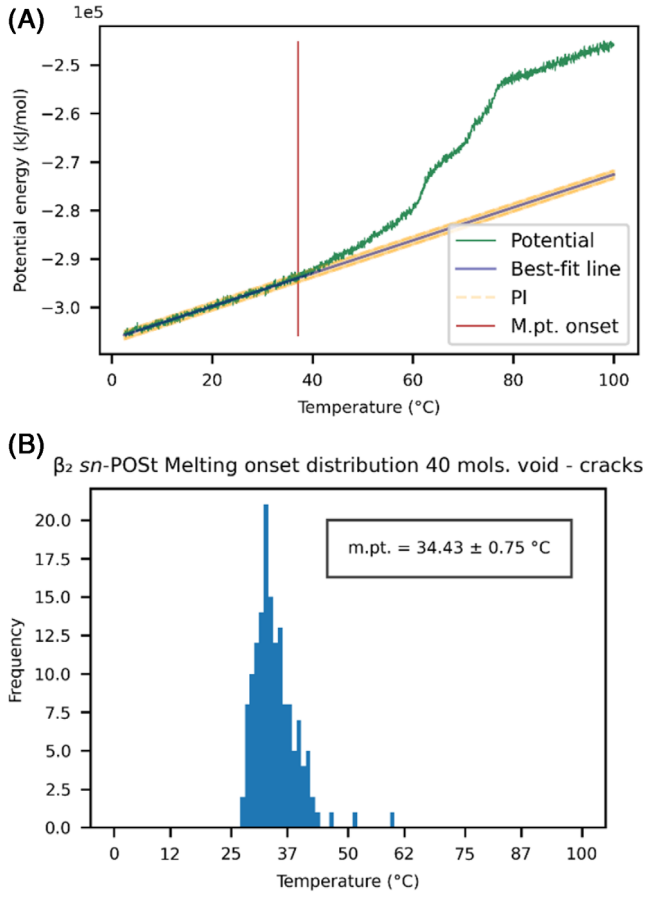

FIGURE 5 (A) Plot of Potential energy (kJ/mol) versus Temperature (°C) for β_2 *sn*-POSt with a 40-molecule crack void (run 4 out of 150). Jagged green line = simulation potential energy as extracted from GROMACS. Solid blue line = Best-fit line as determined by linear regression, using the data points from the initial 5%–30% of the data set. Dashed orange lines/shaded area = Upper and lower 95% Predicted Interval. Red line = Determined melting point onset. (B) Histogram of melting point onset for 150 runs of β_2 *sn*-POSt with a 40-molecule crack void. m.pt. = average melting point onset ±95% confidence interval.

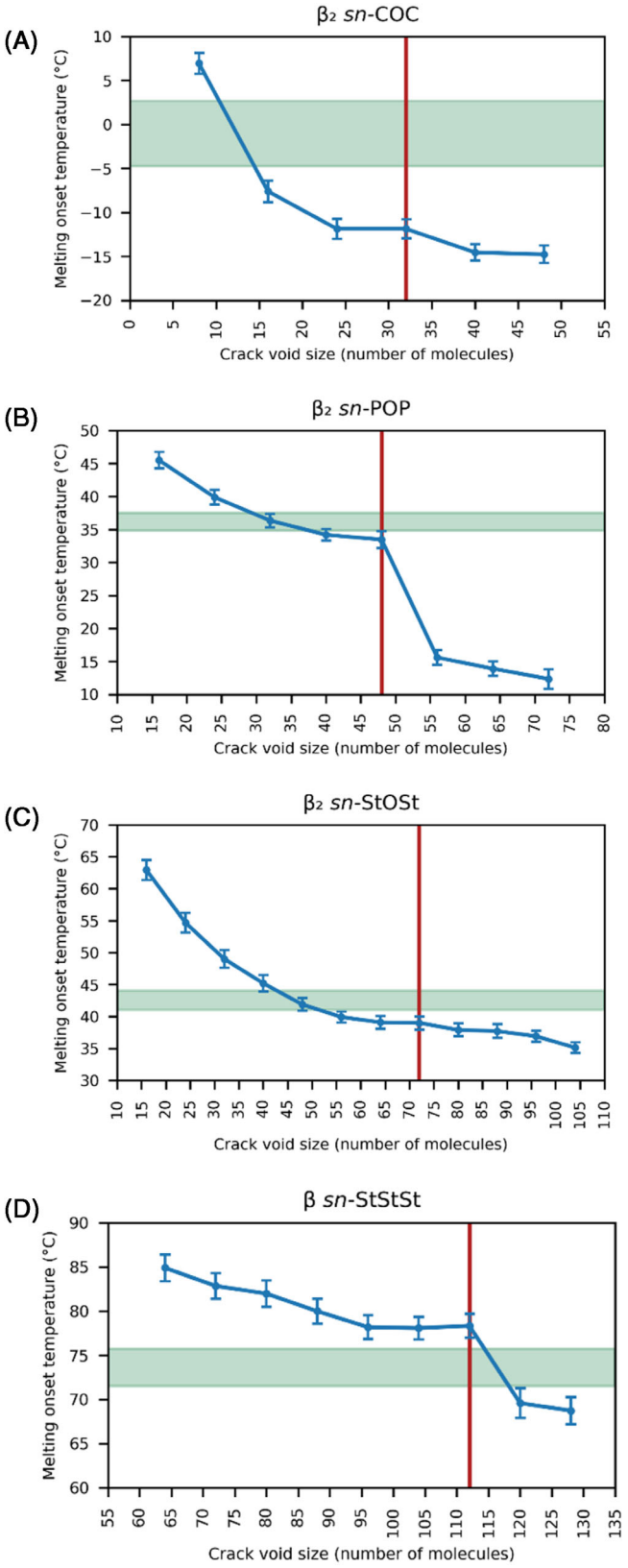

FIGURE 6 Plots of simulated melting temperature (°C) versus void size for (A) β_2 *sn*-COC, (B) β_2 *sn*-POP, (C) β_2 *sn*-POSt, (D) β *sn*-StStSt. Blue markers/line = Melting point onset temperature (°C); error bars = 95% CI; green shaded area = empirical melting point range; red line = melting point for the TAG.

TAG	Empirical melting temperature (°C)	Simulated melting temperature (°C)	Void size at plateau
β_2 *sn*-COC	−4.8[a18]	−11.9 ± 1.1	32
	1.6–2.7[b]		
β_2 *sn*-POSt	33.8–38.0[a13,19–21]	34.4 ± 0.8	40
	35.6–37.6[b]		
β_2 *sn*-POP	34.8–37.5[a21–26]	33.5 ± 1.3	48
	37.3–37.7[b]		
β_2 *sn*-StOSt	41.0–44.0[a21,22,26–29]	38.9 ± 1.0	72
	44.1–44.9[b]		
β *sn*-StStSt	71.5–75.7[a30–35]	78.3 ± 1.4	112
	69.6–72.7[b]		

TABLE 1 Published melting temperatures of various TAGs and simulated melting temperatures ±95% CI.

[a]Empirical values.
[b]Predicted values obtained from the triglyceride property calculator and references therein.[36]

150 simulations. This number of simulations proved enough for all TAGs and void sizes (see Data S5 for all bootstrap error and average melting onset plots, and melting onset distributions).

3.3 | Pure TAG crystals melting point

The determination of the melting point was carried out for five different TAGs, namely β *sn*-StStSt, β_2 *sn*-StOSt, β_2 *sn*-POP, β_2 *sn*-POSt and β_2 *sn*-COC. These TAGs were chosen as they melt at very different temperatures (from 72°C for β *sn*-StStSt to −5°C for β_2 *sn*-COC), are a mixture of fully saturated and monounsaturated TAGs, and β_2 *sn*-StOSt, β_2 *sn*-POP, β_2 *sn*-POSt have similar, but different melting points, with these varying as β_2 *sn*-StOSt > β_2 *sn*-POP > β_2 *sn*-POSt.

Given the fast heating rate of the MD simulations, most crystals did not exhibit a sharp discontinuity when going from solid to melt at any given temperature and instead showed a gradual melting, starting at the void nucleus and extending towards the rest of the crystal, observed as a gradual change in the potential energy (Figure 5A). Given this, the melting point of that crystal is assumed to be the same as the melting onset temperature (red line in Figure 5A), that is, the temperature at which the potential energy of the system started to deviate from the expected (predicted) value (upper orange dashed line in Figure 5A). For each TAG and void size the melting points showed a distribution (Figure 5B), with the melting point determined as the arithmetic mean, and the 95% CI determined from the standard deviation and the number of simulations.

Similar to what was observed for β_2 *sn*-POSt, on increasing the void size for any TAG, a drop in the average melting point was observed until a plateau was reached, followed by another drop due to mechanical collapse (Figures 3 and 6A–D). Comparing the determined melting points with the empirical melting points, these showed very good agreement (Table 1, Figure 7). Not only was the general trend from low to high melting point reproduced, but the determined melting points were also close to the empirically determined values, as well as the predicted values from different models.

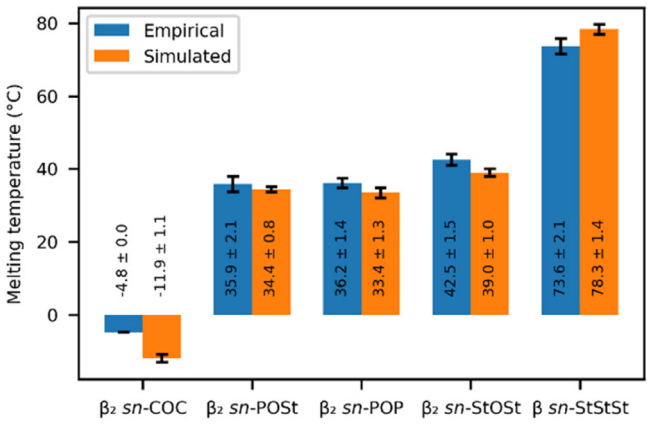

FIGURE 7 Plot of empirical and simulated melting temperatures (°C) versus void size for different TAGs. Error bars = empirical melting point range/simulated 95% CI.

One thing of note from these results is that the plateau was reached at different void sizes for the different TAGs (Table 1). While the exact value will be dependent on the specific void sizes being used to determine where the plateau is reached, a general trend can be observed, namely, the void size required to reach the plateau increases with the melting point of the crystalline TAG. This is due to the increased thermodynamic stability of the crystal, and, hence, the increased need for a lowering of the melting energy barrier. This also means that the ideal void size for any given TAG cannot be known a priori but needs to be determined for each system. Given an initial crystal size of 800 molecules, a void size of 32–112 molecules (at the plateau) means that 4%–14% of the molecules were removed from the perfect crystals to create the defects. As stated by Eike et al.,[2] this level of simulated defect within the crystal is orders of magnitude larger than that observed in real crystals, but this approach nonetheless provides a robust methodology to determine the thermodynamic melting point.

4 | CONCLUSIONS

This study has shown that the thermodynamic melting point of pure TAG crystals can be estimated accurately. This is achieved by introducing void defects into a perfect crystal, heating the resulting structure and determining the melting point onset. Due to the non-equilibrium nature of a melting simulation, an increased number of simulations per void size are required to estimate an accurate melting point, although we find that after 150 simulations the results are sufficiently converged. All simulations were carried out using the COGITO FF,[11] which has proven to be able to reproduce the thermodynamic melting point of various TAGs, despite not being specifically parameterized against this macroscopic property.

ACKNOWLEDGMENTS

The authors thank Mondelēz International for funding this work.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in University of Strathclyde KnowledgeBase at https://doi.org/10.15129/3cbe0feb-386e-4337-b386-4226a34247fe.

ORCID

Tell Tuttle https://orcid.org/0000-0003-2300-8921

REFERENCES

[1] P. M. Agrawal, B. M. Rice, D. L. Thompson, *J. Chem. Phys.* **2003**, *118*, 9680.
[2] D. M. Eike, J. F. Brennecke, E. J. Maginn, *J. Chem. Phys.* **2005**, *122*, 14115.
[3] D. Eike, E. Maginn, *J. Chem. Phys.* **2006**, *124*, 164503.
[4] P. M. Agrawal, B. M. Rice, D. L. Thompson, *J. Chem. Phys.* **2003**, *119*, 9617.
[5] J. F. Lutsko, D. Wolf, S. R. Phillpot, S. Yip, *Phys. Rev. B* **1989**, *40*, 2841.
[6] J. Solca, A. J. Dyson, G. Steinebrunner, B. Kirchner, H. Huber, *J. Chem. Phys.* **1998**, *108*, 4107.
[7] S. L. Zhang, X. Y. Zhang, L. Qi, L. M. Wang, S. H. Zhang, Y. Zhu, R. P. Liu, *Phys. Rev. B* **2011**, *406*, 2637.
[8] A. A. Gavezzotti, *J.Mol. Struct.* **1999**, *485-486*, 485.
[9] N. D. Orekhov, V. V. Stegailov, *High Temp.* **2014**, *52*, 198.
[10] Y. Zou, S. Xiang, C. Dai, *Comput. Mater. Sci.* **2020**, *171*, 109156.
[11] R. J. Cordina, B. Smith, T. Tuttle, *J. Chem. Theory Comput.* **2023**, *19*, 1333.
[12] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.* **2005**, *26*, 1701.
[13] R. Peschar, H. Schenk, J. B. van Mechelen, *Acta Cryst. Sect. B* **2006**, *62*, 1131.
[14] A. Van Langevelde, R. Peschar, H. Schenk, *Acta Cryst. Sect. B* **2001**, *57*, 372.
[15] A. van Langevelde, K. van Malssen, F. Hollander, R. Peschar, H. Schenk, *Acta Cryst. Sect. B* **1999**, *55*, 114.
[16] B. Efron, *Ann. Stat.* **1979**, *7*, 1.
[17] B. Efron, R. Tibshirani, *R. Stat. Sci.* **1986**, *1*, 54.
[18] L. Wesdorp. Ph.D. Thesis. Delft University of Technology (TU Delft). **1990**.
[19] E. S. Lutton, *J. Am. Chem. Soc.* **1951**, *73*, 5595.
[20] S. M. Ghazani, A. G. Marangoni, *Cryst. Growth Des* **2018**, *18*, 4811.
[21] T. Arishima, N. Sagi, H. Mori, K. Sato, *J. Jap. Oil Chem. Soc.* **1995**, *44*, 431.
[22] F. L. Jackson, B. F. Daubert, C. G. King, H. E. Longenecker, *J. Am. Chem. Soc.* **1944**, *66*, 289.
[23] T. Malkin, B. R. Wilson, *J. Chem. Soc.* **1949**, *26*, 369.
[24] H. Lavery, *J. Am. Oil Chem. Soc.* **1958**, *35*, 418.
[25] L. Bayés-García, T. Calvet, M. Àngel Cuevas-Diarte, S. Ueno, K. Sato, *Cryst. Eng. Comm* **2013**, *15*, 302.
[26] S. M. Ghazani, A. G. Marangoni, *Cryst. Growth Des* **2019**, *19*, 704.
[27] L. M. Meara, *J. Chem. Soc.* **1945**, 22.
[28] R. E. Timms, *Prog. Lipid Res.* **1984**, *23*, 1.
[29] K. Sato, T. Arishima, Z. H. Wang, K. Ojima, N. Sagi, H. Mori, *J. Am. Oil Chem. Soc.* **1989**, *66*, 664.
[30] C. E. Clarkson, T. Malkin, *J. Chem. Soc.* **1934**, 666.
[31] E. S. Lutton, *J. Am. Chem. Soc.* **1945**, *67*, 524.
[32] G. H. Charbonnet, W. S. Singleton, *J. Am. Oil Chem. Soc.* **1947**, *24*, 140.
[33] M. Matovic, J. C. van Miltenburg, J. Los, F. G. Gandolfo, E. Flöter, *J. Chem. Eng. Data* **2005**, *50*, 1624.
[34] E. Da Silva, S. Bresson, D. Rousseau, *Chem. Phys. Lipids* **2009**, *157*, 113.
[35] S. M. Ghazani, A. G. Marangoni, *Cryst. Growth Des* **2023**, *23*, 1311.
[36] A. S. Moorthy, R. Liu, G. Mazzanti, L. H. Wesdorp, A. G. Marangoni. Triglyceride Property Calculator. https://lipidlibrary.shinyapps.io/Triglyceride_Property_Calculator/ (accessed: January 2023).

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: R. J. Cordina, B. Smith, T. Tuttle, *J. Comput. Chem.* **2023**, *44*(21), 1795. https://doi.org/10.1002/jcc.27128