# *Ab-initio* prediction of the melting point of organic solids

**A thesis submitted for the degree of Master of Philosophy of Imperial College**

**by Rallia-Iliana Velliou**

**Department of Chemical Engineering**
**Imperial College London**
**London SW7 2AZ, United Kingdom**
**June 2015**

**Declaration of Originality**

I, Rallia-Iliana Velliou, hereby confirm that I am the sole author of the written work enclosed and all sources used to compile it have been clearly referenced.

**Abstract**

The melting point is one of the most fundamental and practically important properties of a compound. For this reason molecular simulation methods have been developed aiming towards accurate computation of melting points. Knowledge of the melting point before a compound has been synthesized could significantly accelerate the design of new materials. Generally, the molecular simulation methods developed so far for the computation of melting points are not fully predictive, since they require an experimental crystal structure as input. An interesting and challenging task is the prediction of the melting point of a compound from first principles- given just the molecular diagram.

In this work, the concept of predicting the melting point of a given organic compound using as an input a computationally obtained crystal structure is investigated. To ensure reliable predictions, it is essential to develop an understanding of how the level of detail of the force fields in terms of crystal structure prediction (CSP) as well in melting point prediction affects the accuracy of the calculations. To explore these requirements the proposed approach in this work combines the application of a CSP multistage methodology [38] developed by the Molecular Systems Engineering group at Imperial College and the freeze method [61] which was recently developed in the group.

Using the proposed approach, two different force fields are employed in this study. Initially, the freeze method is applied to the well known Lennard-Jones potential. Moving on to an organic compound, the case of benzene is investigated. A CSP search is performed and the computational structure is used for the freeze method. Proper choice of force field can affect the agreement with experimental data. For this reason two different force fields are employed in this part of the study, a standard CSP force field and a version of the OPLS force field.

# Contents

# 1 Introduction

The melting point and melting-related properties (e.g. solubility) are important in many fields of research. Although the melting point is the most commonly reported property of organic compounds and often the first property measured after a new compound is synthesized, it is one of the most difficult physical properties to predict [43][48]. Theoretical predictions of the melting point have a long history, and have been based on a wide variety of calculation approaches, as well as different levels of accuracy in their predictions.

The molecular simulation methods developed for the computation of melting points achieve different levels of accuracy in their predictions. What they have in common is that generally most of them are not fully predictive, since they require experimental information on the most stable crystal structure that is adopted by the compound. The accuracy of the computation of the melting point of a given compound from first principles is very important and could accelerate the design of new materials. This idea is investigated in this study, where a technique for performing crystal structure prediction is combined with a melting point prediction method. Accurate prediction depends on the force field applied.

In chapter 2 of this thesis, molecular simulation methods developed so far for the computation of melting points are reviewed. Their general objectives and applications are outlined. In chapter 3, the proposed methodology for this work is presented. The stages of the CSP methodology as well as the freeze method are described. In chapter 4, results are presented. First, the results for the application of the freeze method to the Lennard-Jones potential are reported. Results for the application of our methodology to benzene follow. Specifically, several systems are investigated: two different crystal structures that obtained during crystal structure prediction and match the known structures of benzene, and two force fields. Finally, alternative ways to further examine the effect of the potential used in our methodology are outlined in Section 5.

# 2 Literature Review

For a given pressure, the formal thermodynamic definition of the melting point is the temperature at which the solid phase and the liquid phase of a compound have the same free energy. This means that at the melting temperature the change in Gibbs free energy ($\Delta G$) of the material is zero.

From a thermodynamics point of view, at the melting point the change in Gibbs free energy ($\Delta G$) of the material is zero, but the enthalpy ($H$) and the entropy ($S$) of the material are increasing ($\Delta H$, $\Delta S > 0$). Melting happens when the Gibbs free energy of the liquid becomes lower than that of the solid for that material.

Computational methods developed so far for calculating the melting point can be categorized into two groups: direct methods and "free energy methods" [73][72]. The first group of methods includes the hysteresis method [49][47], the voids method [63][58][50][68] and solid-liquid interface-based methods [53][69]. These methods are based on the direct simulation of the melting process in a dynamical manner. Their application is not very complicated but the accuracy of their results can be limited. The second group of methods includes the Hoover and Rees single-occupancy cell method [32][31], Frenkel and Ladds Einstein crystal method [26] and the $\lambda$-integration method developed by Grochola and co-workers [28][29]. This method was extended by Maginn's group [18][20] and it is also known as the pseudo-supercritical path (PSCP) method. Free energy methods, as their name implies, involve the explicit computation of free energy. These approaches can be more accurate compared to direct methods but their application is generally more complicated and they suffer from high computational costs.

In the following review these methods, their general objectives and applications are presented. First, there is a short description of the direct methods. Specifically, the hysteresis method, the voids method and the interface methods are presented. Focusing on general concepts,the systems they were applied to and the level of accuracy and applicability. A review of the free–energy based methods follows, the thermodynamic integration method, the pseudo–supercritical pathway method and the phase–switch method are included. A brief description of the Quantitative Structure-Property Relationship (QSPR) model methods is also included. Finally, a comparison of the above methods is carried out based on recent reviews and publications.

## 2.1 Direct Methods

As mentioned above, direct methods involve the direct simulation of the melting process in a dynamical manner. In these approaches, an interface between the solid and liquid phases is created and the temperature and pressure that yield a stable interface determines the melting point at that pressure [24][53][42].

Although interfacial methods work reasonably well and their application is quite straightforward, there are concerns about whether finite-size effects are properly accounted for, and different crystal surfaces may have different apparent melting points [73] [72]. Also, most interfacial simulations are run using systems with few to no intramolecular degrees of freedom [4].

In the case of molecular crystals, there is a probability that interfacial MD simulations sufficiently examine the phenomenon of heterogeneous nucleation on MD time scales [5].

### 2.1.1 Hysteresis method

The most straightforward way of finding the melting temperature of a compound is to carry out molecular dynamics simulations of a perfect crystal lattice at increasing temperatures. Although it can be expected that the temperature at which the lattice melts corresponds to the melting point, unfortunately, in molecular simulations this is not the case. When simulating directly a perfect crystal with periodic boundary conditions, only homogenous nucleation melting can occur. This mechanism is comparable to homogeneously nucleated condensation for gases. The free energy barrier for formation of a solid–liquid interface in a perfect crystalline solid causes superheating of the crystal before it melts. This existence of superheating, causes a significant overestimation of the melting point even for simple monatomic molecules. Similarly, when cooling down a liquid, there is an underestimation of the phase transition temperature due to the existence of supercooling. The hysteresis is observed when heating a crystal or cooling a liquid. The hysteresis method was developed based on the homogeneous nucleation melting theory [24]. The thermodynamic melting point depends on superheating and supercooling temperatures as:

$$T_m = T^+ + T^- - \sqrt{T^+ T^-} \tag{1}$$

where $T^+$ and $T^-$ are the observed phase change temperatures when heating a crystal and cooling a liquid, respectively, during a simulation [49][47][47]. The superheating temperature $T^+$ is usually determined easily and accurately from a molecular dynamics simulation, but for complex molecular systems the supercooling temperature $T^-$ is very hard to observe because crystal nucleation does not occur often. It has been found empirically that the supercooling temperature $T^-$ can be equivalent to the glass transition temperature $T_g$. Another complication can arise because the perfect crystals of some solids can be superheated to very high temperatures without melting. For this reason, the hysteresis method has been applied mostly to atomic solids (Lennard–Jones and metals) [48], although nitromethane has also been studied [75].

In this latter work, MD simulations were used to investigate the thermodynamic melting point of crystalline nitromethane, the melting mechanism of superheated crystalline nitromethane, and the physical properties of glassy nitromethane. The maximum superheating and glass transition temperatures of nitromethane were calculated to be 316 and 160 K, respectively. Using the hysteresis method [49] and by taking the glass transition temperature as the supercooling temperature, a value of 251.1 K was calculated for the thermodynamic melting point, which is in good agreement with the two-phase result [2] of 255.5 K and measured value of 244.73 K. Despite this encouraging result, the accuracy of the hysteresis method can be characterized as relatively low [72]. Therefore this method is of limited application.

### 2.1.2 Voids Method

An ideal crystal can melt at a temperature higher than the experimental melting point, because the imperfections in a real crystal such as voids lead to a metastable crystal [58][50]. This observation led to the development of the voids method. In this method, constant pressure molecular dynamics simulations are performed on the crystal phase at increasing temperature [72]. It is expected that when the density changes abruptly, a first order melting transition is observed.

However, this occurs at higher temperatures than the expected melting point. There are cases in which even several hundred degrees of superheat are needed for melting to be enabled. The reasons for that are the small time scales accessible by molecular dynamics simulation and the free energy barrier that must be overcome for the observation of homogeneous nucleation

in a perfect crystal. The nucleation free energy barrier is lowered with the creation of voids in the crystal which cause melting to begin on the time- and length-scales accessible by molecular dynamics simulation. The voids can be introduced by removing molecules or ion-pairs in the case of ionic solids. The observed melting point reduces while the void density increases and it has been observed that as the void density continues to increase, there comes a point when the melting point does not relate any more to the void density. Specifically, it has been observed that when the void density is between 6% and 10%, the melting point levels off and is considered to be the actual melting point. When the void density is above 10, the crystal is no longer mechanically stable. The voids method can be applied a sufficient amount of the solid remains in the simulation supercell to maintain the integrity of the crystal structure. In the case of too many voids being distributed in the cell, the solid becomes mechanically unstable and collapses without a discontinuous solid-liquid phase transition [5].

The voids method has been successfully applied to a number of materials that include simple systems such as rare gas solids [1][63] as well as complex molecular and ionic crystals such as nitromethane [2], ammonium nitrate [68] and l-ethyl-3-methylimidazolium hexafluorophosphate [4]. The drawbacks of this method are that it requires a lot of computations because it is necessary to determine the melting point for several void densities. In some cases the plateau region is narrow and ambiguous, requiring more simulations.

Alavi and Thompson [5][4] used this method to study the liquid-state properties and melting of 1–ethyl–3 methylimidazolium hexafluorosphosphate [emim][PF6]. The melting point was determined by equilibrating the solid-state supercells in which void defects had been introduced to eliminate the free-energy barrier for the formation of a solid-liquid interface. The computed melting point was 375 K, which is approximately 12% higher than the experimental value of 331 K.

Generally, for simple systems the voids method is reasonably accurate, but in the case of more complex molecules the application of the method is complicated and the calculation [5][4].

### 2.1.3 Interface Methods

A way of lowering the nucleation free energy barrier is to simulate a two-phase system with a solid–liquid interface [53]. As mentioned above, this lowering of the nucleation free energy barrier allows melting to occur on the

time and length scales accessible to molecular dynamics simulation. Based on this idea at least two methods have been introduced [69] and will be briefly described.

The first, considers as the melting point the temperature at which certain system properties such as density, appear to have some discontinuity similarly to the methods mentioned above, but in this case the nucleation barrier is lowered by the solid-liquid heterogeneous interface. Constant pressure and constant temperature curves are generated at different temperatures.

The second method involves the creation of a solid and liquid interface which equilibrates at a given temperature and volume ($NVT$), continuing with simulations in the $NVE$ ensemble. After equilibration, the simulation box size is changed, the density is perturbed and as a result the total energy of the system is either decreased or increased. Then another $NVE$ simulation is performed and as a result either a part of the solid will start melting or a part of the liquid will crystallize, redistributing the potential energy and kinetic energy until a new equilibrium is reached. The liquid-solid interface will continue, if the perturbation is small. The system's average temperature and pressure are taken giving one point on the solid-liquid coexistence curve [72]. The procedure is repeated until a set of equilibrium pressure/temperature points are obtained. The melting point of the compound at a certain pressure is the temperature on the curve for that pressure.

The interface methods have been applied to a wide range of materials such as Lennard–Jones fluids[1], MgSiO3 [8], NaCl and MgO [10], LiF [9], fcc and hcp metals [46], silicon [71], and nitromethane [2]. However, these methods require more than 500 molecules and long simulation times for accurately determining the equilibrium melting point, therefore it is difficult to apply to the crystals of large molecules [4].

In summary, both methods may give accurate prediction of the melting point in the case of simple molecules, but for more complex systems fails to give a reliable estimation.

## 2.2 Free-energy based methods

The second major category of melting point simulation methods is based on satisfying the phase equilibrium conditions, specifically equality of temperature, pressure, and chemical potential [72]. It includes a number of techniques, such as thermodynamic integration [27] and phase switch [70] approaches. Although free-energy based methods provide a rigorous way of

computing the melting point of a given compound, they are generally more complicated to apply compared to the direct methods described in the previous section.

### 2.2.1 Thermodynamic Integration

Thermodynamically rigorous pathways for predicting the melting point of solids have been developed, which require calculating the free energy of the solid and liquid phases, and include thermodynamic integration based on the Kirkwood coupling parameter method [39]. These methods involve simulating a single phase at each step of the procedure and therefore avoiding the complication of interfaces in the simulation. When simulating two-phase systems it is important that the number of molecules in the interfacial region is not a large fraction of the total number of molecules in the simulation [4][25]. To calculate the free energy difference between states I and II, the potential energy of the system U is written as the sum of the potential energies of the two states, coupled by the parameter $\lambda$:

$$U(\lambda) = (1 - \lambda)U_I + \lambda U_{II} \tag{2}$$

State I can represent the system of interest and state II can be a reference state with a related potential energy function. The potential energy corresponds to that of state I for $\lambda$=0 and state II for $\lambda$=l [5].

As mentioned above, the free energy methods are based on satisfying the phase equilibrium conditions and they tend to avoid nucleation phenomena by explicitly computing the free energy [33]. Here, as in direct methods, hysteresis is also an issue since it can cause errors related to the first-order phase transition between the liquid and crystal phases and has to be taken into consideration while designing these methods. In most cases thermodynamic integration is applied for calculation of the free energy change along a carefully designed path. Two methods developed for computing melting points based on thermodynamic integration are the single-occupancy cell method of Hoover and Ree [32] and the Frenkel and Ladds Einstein crystal method [26].

The single-occupancy cell method of Hoover and Ree [32][31] involves a particle system which is connected to a low density lattice. Each particle occupies an individual cell and the free energy is known analytically. When the density is high enough, the interaction between the cell walls and particles

are minor and the real system is restored. Hysteresis is avoided by controlling the solid-liquid phase change to occur at low density or not at all.

Similarly to the single-occupancy cell method, the Frenkel and Ladds Einstein crystal method [26] involves the connection through a reversible path of the solid phase to an Einstein crystal with the same crystallographic structure, the free energy of which is known analytically. This reference state can be reached from the initial solid by slowly switching on harmonic springs which couple the atoms to their lattice sites. As the Einstein crystal has an identical structure to the initial solid, it is most likely that this path between the two will be free of phase transitions and therefore reversible. Similarly, the liquid phase is coupled to an ideal gas or some other state, for which the free energy is also known. Once the free energy differences between the solid and liquid and the reference states are known, the free energy difference between the actual solid and liquid can then be computed. This method is fast and accurate, and can be applied for cases of molecular solids and solids containing defects. Though theoretically rigorous, this method is difficult to apply to complex molecular crystals and computationally demanding[5][75].

### 2.2.2 Pseudo-supercritical pathway method

Another free energy-based procedure for rigorously computing the melting point is the pseudo-supercritical pathway method (PSCP method). It has been suggested recently and it links directly the solid and liquid phases through a multi-stage pseudosupercritical pathway which avoids the discontinuous phase transition. It has been recently developed by the group of Maginn by extending a thermodynamic integration technique [19][21] originally proposed by Grochola [28][29]. It is an inter-phase approach [28][18] in which the solid and liquid phases coexist and are connected directly by intermediate states so analytical reference states are not necessary. This method involves the use of a pseudosupercritical transformation path to transform reversiblybetween solid and liquid phases. Integration along this path yields the free energy difference at a single state point, which can then be used to determine the free energy difference as a function of temperature and therefore locate the coexistence temperature at a fixed pressure.

The intermolecular interactions are scaled down in the intermediate states and the phase change is controlled to occur in a reversible way so that the error caused by superheating is minimized. This method has been applied to a number of molecular systems of varying complexity [18]. The application

of the method includes two steps. During the first step, isothermal-isobaric $NPT$ MD simulations are run for the liquid and crystalline stages in order to obtain relative free energy curves. The enthalpy of each pure phase is computed as a function of temperature, and the Gibbs Helmholtz equation is integrated to obtain Gibbs free energies relative to an arbitrary reference temperature $T_{ref}$.

During the second step, the free energy difference between the pure liquid and pure crystalline phase along a thermodynamic path is calculated using thermodynamic integration. From this, the temperature at which the liquid and crystalline free energies are equal can be found, which corresponds to the melting temperature. This procedure differs from commonly used reference state approaches [27], where the absolute free energies of the liquid and crystal are obtained by integrating along two separate paths to reference states of known free energy. There are five separate states along the thermodynamic path, all at the reference temperature:

1. the liquid at a density corresponding to the pressure of interest
2. a weakly interacting fluid at the liquid density
3. a weakly interacting fluid at the crystalline phase density
4. an ordered weakly interacting state at the crystal density
5. the crystal at the pressure of interest

The transitions between the states (except states 2 and 3) are carried out by changes in a coupling parameter $\lambda$ and the Helmholtz free energy change between states is given by:

$$\Delta A_{i \to j} = \int_0^1 -\langle \frac{\partial U}{\partial \lambda} \rangle d\lambda \tag{3}$$

The free energy change between states 2 and 3 is given by the thermodynamic identity [33]:

$$\Delta A_{2 \to 3} = \int_{V^l}^{V^S} -\langle P \rangle dV \tag{4}$$

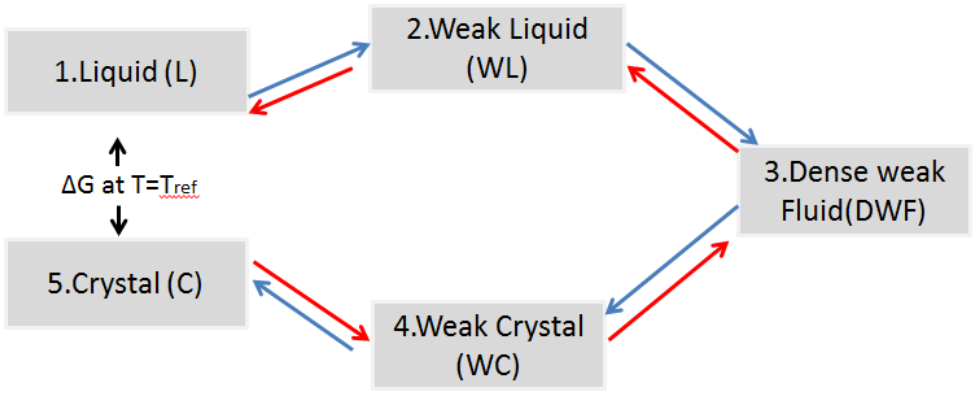


Figure 1: A schematic of the pseudo-supercritical path for melting point calculations [72].

The PSCP method [19][21] has been applied to simple cases such as the Lennard-Jones fluid [19] and to more complex molecules such as to molten sodium chloride and a number of ionic liquids. It has also been extended to deal with multiatom molecules such as benzene and triazole [21]. The method, which is referred to as constrained fluid integration or pseudosupercritical path sampling, overcomes some of the limitations of traditional thermodynamic integration methods because it does not require fluid and solid reference state free energies.

### 2.2.3 Phase-switch method

Another example of free-energy based methods for computing melting points is the phase switch method [70]. This approach can be considered to be similar to the PSCP method, since is based on the transformation between the solid and liquid states through gateway states. The free energy is calculated as a ratio of probabilities of the frequency of visits made to the two states. The method has been applied for cases of simple systems giving accurate predictions, but in the cases of complex multi-atom molecular systems this method has still not been used. Orkoulas and co-workers have used a similar method based on an extension of the single-occupancy cell method [56][55].

## 2.3 QSPR methods

Another group of computational methods for melting point prediction is that of the Quantitative Structure-Property Relationship (QSPR) model methods [57][40][23] . The development of these methods comes as a result of the attempts to create a quantitative link between melting point and structure [37][66]. QSPR has been used to calculate melting points for a variety of systems including ionic liquids. Generally, the application of QSPR methods is quite simple. In order to parameterize the model a set of experimental data is required [40]. Therefore, the accuracy of the method relies on the availability of the experimental data set for a certain compound and it is usually limited. QSPR has been applied for melting point prediction for a number of systems [37]. The parameters are generally valid only for compounds similar to those in the training set, which makes these models less useful when new compounds are being developed. However,the main disadvantage is that in order for a prediction to be performed, for a given compound a fairly large amount of experimental data of similar compounds is required [57][72].

## 2.4 Freeze Method

A melting point prediction method was recently developed in the Molecular Systems Engineering group involving the use of a direct interfacial methodology to evaluate solid-liquid equilibrium. The freeze method [61] determines a solid-liquid coexistence point and is carried out in three main steps. Initially, solid and liquid phase isobars are calculated to determine the hysteresis region. This is achieved by heating a crystalline solid at constant pressure in the $N\sigma T$ ensemble until melting occurs and the system is then cooled at constant pressure until it refreezes to a solid. Then, a solid-liquid coexistence system is created and is relaxed under the $NVT$ ensemble. If solid-liquid coexistence is maintained the coexistence ($T$,$P$) conditions are used as initial points for Gibbs-Duhem thermodynamic integration [41] according to the Clausius-Clapeyron formula. The freeze method allows the simulation of two coexisting phases avoiding complications observed in other simulation techniques. The technique was initially tested for the LJ potential for which there is appreciable data in the literature for comparisons to be made. A detailed outline of the method follows in the Methodology section.

## 2.5 Comparison of methods

In the past years there has been a lot of activity in the area of melting point prediction methods. Despite that, accurate and efficient prediction of melting points for complex molecules is still a challenging task for molecular simulation. Most of the methods developed so far have been validated for relatively simple systems such as the Lennard-Jones fluid. It is necessary to examine how these different approaches perform when they are applied for predicting the melting point of more complex molecules.

Recently, a number of direct methods for melting point prediction were applied and compared by Feng et al. [24]. In this work, molecular dynamics simulations where carried out to investigate the solid-liquid transition of the ionic liquid 1-ethyl-3-methyl imidazolium bromide ([emim]Br) by using the direct heating method, the hysteresis method, the voids method, the sandwich method and the interface/$NVE$ method. In the sandwich method [30][24], solid-liquid interfaces are introduced to eliminate superheating. The liquid phase is in contact with two different planes of the solid phase and when the temperature increases, it is found that the solid slab shrinks slowly, the liquid slab grows slowly, and finally the whole system becomes uniform in the liquid state. The melting points obtained from the first three methods are 547± 8 K, 429± 8 K, and 370± 6 K, respectively; while for the sandwich method, the melting points are 403± 4 K when merging along the x axis by anisotropic isothermal-isobaric ($NPT$) ensemble, 393± 4 K when along the y axis by anisotropic $NPT$ ensemble, and 375± 4 K when along the y-axis by isotropic $NPT$ ensemble. For the interface/$NVE$ method method, when the slabs are merging along different directions (x-axis, y-axis, and z-axis), the melting points are 364± 3 K, 365± 3 K, and 367± 3 K, respectively. The melting points obtained from different methods are approximately 55.4%, 21.9%, 5.1%, 14.5%, 11.6%, 6.5%, 3.4%, 3.7%, and 4.3% higher than the experimental value, which is 352 K. Results and comparison with the experimental value are presented in Table 1.

From the investigation, the voids method and the interface/$NVE$ method were suggested as favourable approaches.

However, in a recent paper on melting point calculation methods comparison, Zhang and Maginn [72] report that it is possible that these two methods were not properly applied in the study. Regarding the voids method, only a single simulation with 16 voids was performed. It is known from previous studies that a series of simulations at different void densities should be carried

Table 1: The melting point ($T_m$) of [emim]Br determined by different simulation methods and experiment in the work of Zhou and Maginn [24].

| Method | $T_m$ | comparison with experimental value |
|---|---|---|
| Direct Heating | 547 ± 8 | 55.4% |
| Hysteresis | 429 ± 8 | 21.9% |
| Voids | 403 ± 4 | 14.5% |
| | 370 ± 6 | 5.1% |
| Sandwich | 393 ± 4 | 11.6% |
| | 375 ± 4 | 6.5% |
| | 364 ± 3 | 3.4% |
| NVE | 365 ± 3 | 3.7% |
| | 367 ± 3 | 4.3% |
| Experiment | 352 | |

out because the resulting melting point depends on the void density. Thus using a single arbitrary void density is not enough for an accurate prediction. For the interface/$NVE$ simulation, Zhou and Maginn state that the initial point of the method was a configuration equilibrated under the $NPT$ ensemble, although only a single NVE trajectory was ran after perturbation. The final temperature obtained was considered to be the melting point. Zhang and Maginn report that since both pressure and temperature are coupled in the $NVE$ ensemble, the equilibrium temperature observed from the $NVE$ simulation corresponds to the average pressure during the simulation, which was not reported by Zhou and Maginn. In their work, Zhang and Maginn apply four melting point computational methods, their free energy-based method (the pseudo supercritical path (PSCP) method) and three direct methods (two interface-based methods and the voids method) for the cases of argon and liquid 1-n-butyl-3 methylimidazolium chloride ([BMIM][Cl]). The performance of each method was compared systematically. The study showed that for the case of argon all the methods applied reproduce the experimental melting point reasonably accurately. For [BMIM][Cl], the melting point was computed to be 320 K using a revised PSCP procedure, which agrees well with the experimental value 337–339 K. However, large errors were observed in the computed results using the direct methods, suggesting that these methods are inappropriate for large molecules with sluggish dynamics.

In another paper by Alavi and Thompson [5] several methods for MD sim-

ulations of melting and the calculation of melting points have been reviewed. The thermodynamic melting point is defined by the equality of the free energies (chemical potentials) of the solid and liquid phase, and this can be used to compute the melting point with MD simulations. Also, the melting point can be determined in simulations corresponding to one of the actual melting mechanisms, namely, homogeneous nucleation melting, surface induced melting, or void (imperfection) induced melting. The free energy determination of the melting point is theoretically rigorous, but the other methods have been shown to have reasonable theoretical justifications, and they are easier to implement. These methods are complimentary and have been used to determine the melting points of a wide range of solids. Void-induced melting simulations allow a straightforward use of periodic boundary conditions at different pressures in $NPT$ simulations and avoids problems associated with superheating [72]. The simulations are straightforward to set up and require between 100 and 200 molecular or ion pairs to obtain converged results. More than 800 molecules can be required for a two-phase $NVE$ solid–liquid simulation.

For cases of molecular and ionic salts composed of small, relatively rigid molecules and ions, potential energy functions determined to reproduce solid- or liquid-state properties (far from the phase transition) appear to predict accurate melting points. Also, standard force fields such as AMBER have also been shown to reproduce the melting point with good accuracy for these kinds of solids. For large floppy molecules with internal degrees of freedom active at temperatures near the melting point, it is much more critical that the force field accurately describe the active internal modes; quantum chemistry calculations of isolated molecules have proven useful in determining the force constants and barrier to internal rotations.

## 2.6 Conclusions

Although there has been a lot of activity and progress in the area, predicting melting points using computational methods is still an extremely challenging problem [72]. In addition to that, most of the molecular simulation methods developed so far for the prediction of melting points require an experimental crystal structure as input, which means that such calculations are not fully predictive [73]. On the other hand, the prediction of crystal structures (CSP) is a rapidly growing area of research and significant progress has been made. However, many challenges still exist such as structure searching and global minimization methodologies, development of intermolecular model potentials, force fields for molecular flexibility and first principles electronic structure mechanical methods and development of dynamical simulations [15]. It would be very useful to predict the melting point and other physical properties of a given compound by using a predicted crystal structure from a CSP search as a starting point.

# 3 Methodology

Generally, molecular simulation methods developed so far for the computation of melting points are not fully predictive, since they require an experimental crystal structure as input. However, knowledge of the melting point before a compound has been synthesized could significantly accelerate the design of new materials. An interesting and challenging task is the prediction of the melting point of a compound from first principles- given just the molecular diagram. This idea is investigated in this work, where the prediction of the melting point of an organic compound is attempted by using a computationally obtained crystal structure.

In this work, the concept of predicting the melting point of a given organic compound using as an input a computationally obtained crystal structure is investigated. To achieve this a Crystal Structure Prediction (CSP) methodology is applied combined with a melting point prediction method. For obtaining a computational structure a CSP multistage methodology [38] developed by the Molecular Systems Engineering group at Imperial College is applied and the basic stges of the procedure are described. As for the melting point computation, the freeze method [61] which was recently developed in the Molecular Systems Engineering group is used and is outlined. Accurate prediction depends on the force field applied and the melting point computation method. Proper choice of force field can affect the agreement with experimental data. For this reason two different force fields are employed in this study. More details regarding the parameters of the force fields are presented in the Results section.

## 3.1 Crystal Structure Prediction

Crystal Structure Prediction (CSP) is a set of techniques to identify through computation the likely crystal structure of a given compound, usually by applying optimization algorithms in which a lattice energy function is used as an objective function [15][44]. A characteristic of crystal structure prediction is that the interest is not only to locate a global minimum structure, but all low-energy minima. This is due to the importance of metastable polymorphs that are higher in energy than the most stable ones. Also, the search for multiple structures allows to overcome the limitations of existing models, in which model uncertainty may lead to inaccuracies in the energies of different structures.

From a practical perspective, it is known [45][54][17][16][7] that the reliable modelling of an organic crystal requires high accuracy and hence computationally demanding models in order to achieve meaningful results. The Molecular Systems Engineering group at Imperial College has been performing crystal structure prediction studies applying a multistage methodology, which deals with different issues at different stages [38]. These stages are described here in the context of rigid molecules [36]:

- Stage 1: choice of computational model

The first stage of the applied methodology involves the choice of an appropriate computational model. A conformational analysis is performed based on a survey of the Cambridge Structural Database (CSD), and isolated molecule quantum mechanical calculations. For this purpose, a level of theory and basis set are selected for the isolated-molecule QM calculations used to compute and characterize the intermolecular electrostatic interactions. A semi-empirical model is used to represent the dispersion/repulsion interactions with an associated set of parameters. This analysis is used to identify the most accurate combination of level of theory and basis set by comparing to experimental data available.

- Stage 2: global search

During the next stage of the methodology, a global search using the CrystalPredictor [36] algorithm is carried out, in order to identify all possible low energy minima of the lattice energy. The search is carried out over the unit cell parameters and the molecular positions. The number of identified minima is of the order of millions and a simple model of electrostatics (point charges) is adopted to make the computations tractable. Usually, the most promising minima generated during this step are then further minimized using a much more accurate model, in stage 3.

- Stage 3: local search

The most promising minima, which are those with the lowest lattice energy, identified in the previous stage are further minimized using a much more accurate [17], and therefore computationally-demanding model which is based on a distributed multipole expansion of the electrostatic potential [64]. The DMACRYS software is used for this purpose.

## 3.2 Melting Point Prediction

Once a structure is determined by CSP, a melting point (MP) computation follows. Accurate prediction depends on the melting point calculation method and the force field applied. In this work, a method recently developed in the Molecular Systems Engineering group, the freeze method [61], which can be used to determine a solid-liquid coexistence point, is applied. To use this method on real molecules it is very important to apply an appropriate force field. For this reason two force fields are investigated: the force field used in *CrystalPredictor* and the rigid version of the OPLS force field [12]. The freeze method allows the simulation of two coexisting phases avoiding complications observed in other simulation techniques. The method is carried out in three steps as shown in Figure 2:

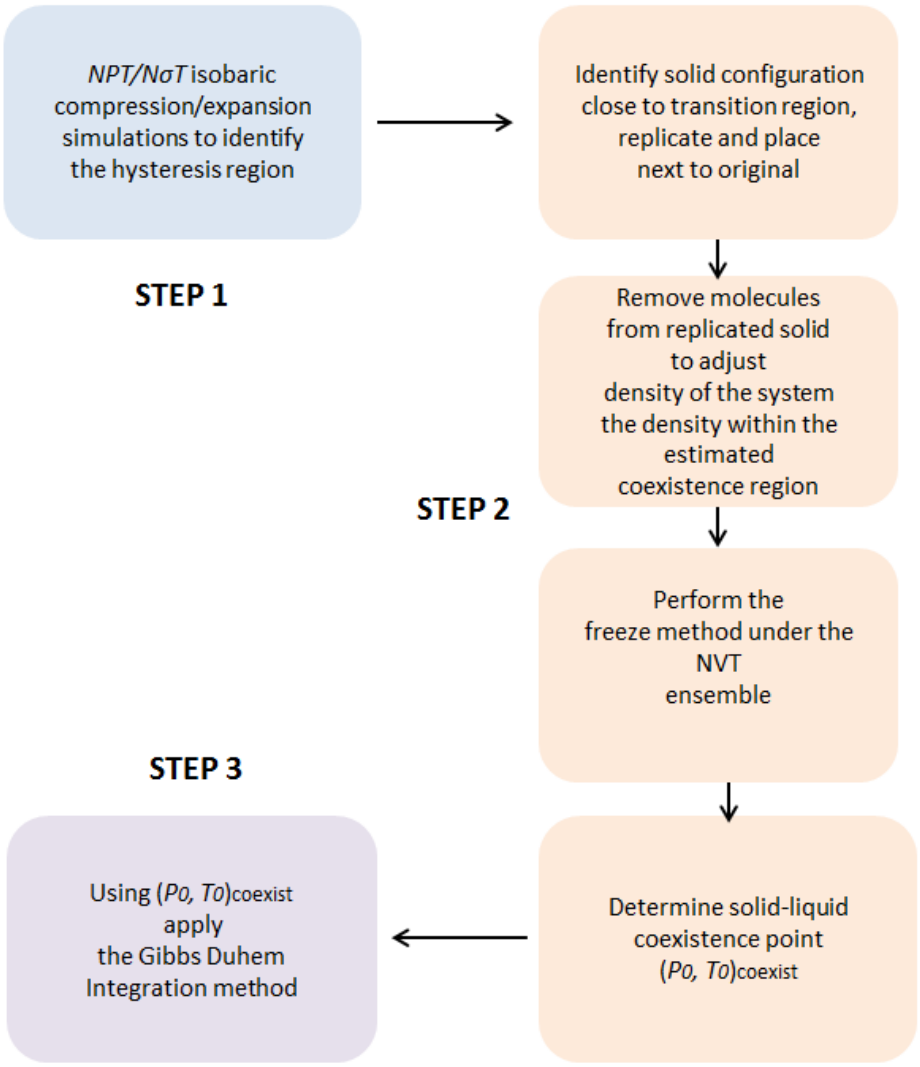


Figure 2: A schematic showing the steps implemented in the freeze method. Each step is highlighted by a different coloured box for clarity.

- Step 1: Solid and liquid phase isobars

In the first step of the method, the hysteresis region of the phase transition must be defined. For this, a crystalline solid configuration is necessary. The initial crystal configuration used in the freeze method is based on a super cell built from the predicted unit cells at 0K and 0 Pa. The unit cell obtained from a CSP search is replicated towards the x,y,z directions resulting a solid configuration.
The resulting crystalline solid configuration is heated at constant pressure in the $N\sigma T$ ensemble in order to allow the shape of the simulation cell to fluctuate. The heating isobar of the solid continues until melting occurs. The system which is now in the liquid phase is cooled down at constant pressure. The solid simulations for the heating isobar are run under the $N\sigma T$ in order to allow the shape of the simulation cell to fluctuate. The simulations of the liquid phase for the cooling isobar are run under the $NPT$ ensemble. Each simulated solid along this isobar represents a thermodynamically stable solid configuration at fixed pressure and temperature. To determine the phases of each state point along the isobar, the radial distribution function of each simulation is analysed. Freezing/melting transitions are first order and are associated with a strong degree of hysteresis, therefore the exact melting temperature can not be obtained via this route. Hysteresis can be attributed to factors such as system-size effects or high free-energy barriers associated with these transitions, which lead to metastable state points and the inability to obtain an accurate transition temperature. Although the exact melting temperature cannot be obtained, the advantage of this method is that the expected region of the phase transition is narrowed to within the boundaries of the hysteresis region which can then be investigated further.

- Step 2: Determining a solid-liquid coexistence point

During the second step of the methodology, the freeze method is applied. The initial solid required for the freeze methodology is obtained from the heating isobar used to determine the hysteresis region. Of these heated solids, the one closest to the melting transition, (i.e. the solid obtained just before the phase transition to the liquid phase occurs) is considered to be closest to the expected phase coexistence region and due to hysteresis this corresponds to a metastable solid which is most likely at a density just below the saturated solid density. Therefore it is this solid configuration which is implemented in the freeze methodology.

Once a solid configuration is selected, the system is replicated four times towards the z-direction. Half of this larger system is fixed frozen. This means that molecules within this part of the box remain in position such that their presence has no effect on the simulation of the other unfrozen half of the simulation cell. In the adjoining box, molecules are removed in order to adjust the overall system density to that at which coexistence is expected. The estimated fixed density of the system is the average of the upper and lower densities ($\varrho_{est} = \frac{\varrho_{upp}+\varrho_{low}}{2}$) and a temperature within the hysteresis region is taken as the fixed temperature of the simulation. As the actual coexistence region is yet to be determined, the temperature is averaged as the midpoint of the hysteresis. The actual value used is not necessary as the next step in the procedure will allow the system pressure to adjust.

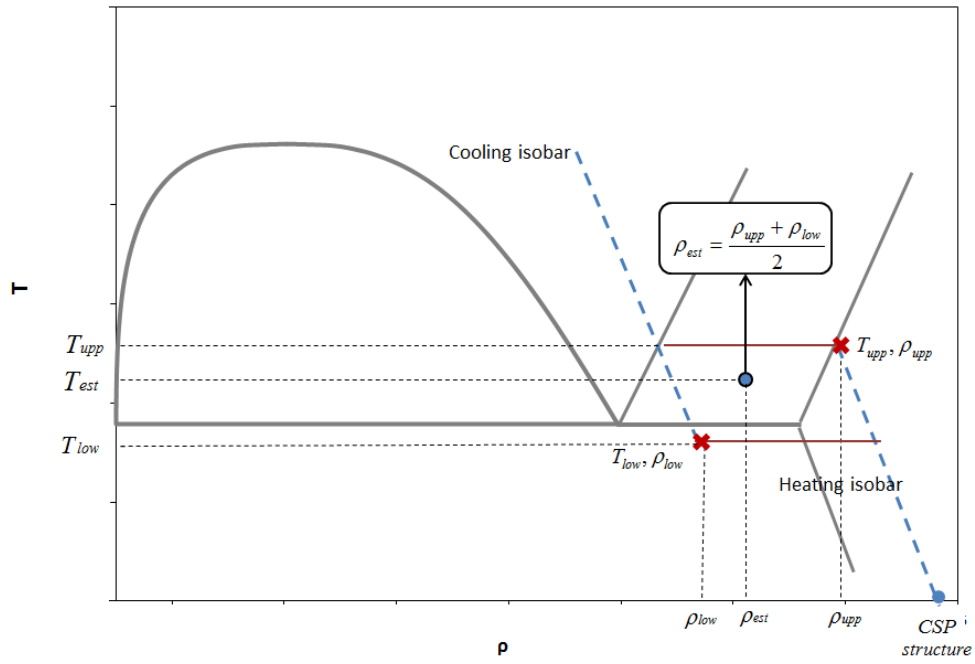


Figure 3: Schematic showing how the estimated conditions $T$ and $\varrho$ are chosen for the freeze method to be carried out. The dashed blue lines are the liquid and solid isobars, the red crosses show the highest temperature simulation along the heating isobar before the solid melts ($T_{upp}$, $\varrho_{upp}$), a lower temperature simulation along the cooling isobar ($T_{low}$, $\varrho_{low}$) and the red lines show the region of hysteresis. The blue closed circle is the system at a $T$ within the hysteresis region and est used in the freeze method. The blue open point corresponds to the initial crystalline solid created from the CSP structure.

The entire system is then relaxed under the $NVT$ ensemble. During this simulation, the unfrozen half is expected to melt into a liquid while the frozen particles remain stationary. Finally, the stationary molecules are released (i.e. all molecules within the box take part in the simulation) and the entire system is allowed to equilibrate under the canonical $NVT$ ensemble. The expected density profile of a successful simulation shows two well defined plateaus corresponding to the coexisting phases. For the case the two coexisting phases are not equilibrated (stable) in the first iteration of the

freeze method, the choice of the metastable density is adjusted to accommodate for this. For the case where the system solidifies, more molecules are removed in the third stage of the algorithm, conversely if the system melts, fewer molecules are removed. This shift in the estimated metastable density is repeated until coexistence is obtained as shown in Figure 4. The densities of each phase are obtained once coexistence is achieved.

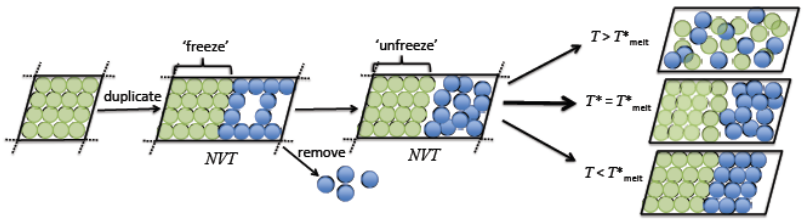


Figure 4: Schematic of the freeze method. The solid black lines define the simulation cell and the extended dashed lines represent the periodic boundary conditions. The green and blue particles are the same, the colour is changed to differentiate between the replicated (blue) and original (green) system.

The freeze methodology requires a previous knowledge of an approximate transition point. Thus, the isobars calculated in the previous step provide estimates of lower and upper temperature boundaries corresponding to the phase transitions upon cooling a liquid and heating a solid respectively. The outcome of the procedure is independent of the density of the initial solid structure chosen as long as it is a relaxed solid (without stress) close to melting. During the final equilibration step when it is in contact with a liquid the interface will rearrange (melt or freeze) to accommodate the corresponding equilibrium densities.

The coexistence pressure is obtained as an output of the canonical $NVT$ ensemble simulation. As the simulation cell is non-cubic and the interface exists perpendicular to the z-axis, the pressure is obtained as the normal component of the pressure tensor i.e. $P_{zz}$. To ensure the accuracy of the output pressures of these simulations an additional simulation is run in the canonical ensemble. As a next step, an independent simulation is run under the $NVT$ ensemble at the coexistence temperature and densities using the density of the solid region obtained from the density profile determined by

the previous coexistence $NVT$ simulation. The output pressure obtained from this single phase simulation corresponds to the coexistence pressure at the fixed volume and temperature conditions. This additional simulation has proven (at least for the LJ sphere) to be slightly more accurate than the coexistence pressure obtained from the normal component of the pressure tensor due to the removal of interface the effect.

- Step 3: Gibbs Duhem Integration (GDI)

Finally, with this known coexistence point the Gibbs-Duhem integration technique can now be implemented to complete the coexistence boundaries for the solid-liquid region. The integration advances from the state of the determined coexistence point according to the Clapeyron formula, a first-order ordinary differential equation that prescribes how the pressure must change with temperature to maintain coexistence [40][67]:

$$\left(\frac{dP}{d\beta}\right)_{\sigma} = -\frac{\Delta h_{melting}}{\beta \Delta V_{melting}} \tag{5}$$

where $\beta = \frac{1}{k_B T}$ and $\Delta H$ and $\Delta V$ are the enthalpy and volume differences per particle respectively between the liquid and solid phases and the differentiation is taken along the saturation line $\sigma$. Values for $\Delta H$ and $\Delta V$ can be obtained from molecular simulation, and the derivative equation can be solved stepwise by numerical integration.
A typical integration step along the saturation line as described in the work of Kofke [40] proceeds as follows: Given that an initial coexistence point $(P_0, T_0)$ is known, a temperature increment $\Delta\beta$ is chosen ($\Delta\beta = \frac{1}{k_B T_1} - \frac{1}{k_B T_0}$) and the saturation pressure $(P_1)$ at the new temperature $(T_1)$ is predicted using the calculated $\Delta H$ and $\Delta V$ at $(P_0, T_0)$. Simulation data is used to solve the right hand side of the Claussius-Clapyeron equation and a new pressure $(P_1)$ is predicted:

$$P_1 = P_0 + \Delta\beta \frac{\Delta h_{melting}(P_0, T_0)}{\beta_0 \Delta V_{melting}(P_0, T_0)} \tag{6}$$

Simultaneous but independent $NPT$ simulations of the coexisting phases are carried out at the predicted conditions $(P_1, T_1)$. The initial configuration for each phase is taken from a previously equilibrated system from the previous step i.e $(P_0, T_0)$. $NPT$ simulations determine the new enthalpies

($\Delta h_{melting}(P_1, T_1)$) and volumes ($\Delta V_{melting}(P_1, T_1)$) necessary for the calculations. Averages taken throughout the simulations are used to correct the estimate of the pressure to convergence. Thus strictly the pressure is not fixed during the simulation. The outcome of the procedure is the complete melting line of the target substance.

# 4 Results

## 4.1 Lennard-Jones

Initially, the freeze method described in the methodology section is applied for the well-known Lennard-Jones model, in a variety of temperature-pressure combinations. All properties in the application of this model are represented in reduced units according to:

$$T^* = T\frac{k_b}{\epsilon} \tag{7}$$

$$P^* = P\frac{\sigma^3}{\epsilon} \tag{8}$$

and

$$\rho^* = \rho\sigma^3 \tag{9}$$

In order to identify the expected solid-liquid coexistence region, extensive isothermal-isobaric simulations are performed in the temperature range of $T^* = 0.675 - 2.7$. The pressure of each isobar is selected by creating a solid system of a density that is in the stable solid phase and performing a simulation under the $NVT$ ensemble. The output pressure obtained through the virial route and final configuration of this simulation are then used as the fixed pressure of the isobar and the starting configuration for the heating isobar.

For the case of the LJ sphere, all simulations are performed using the DLPOLY software for Molecular Dynamics. For determining the hysteresis region and calculating pure-phase properties, simulation s are conducted under the $N\sigma T$ and $NPT$ ensembles. The Nosé-Hoover thermostat and barostat are implemented to ensure that an average constant temperature and pressure are maintained throughout the timescale of each simulation. The system size is chosen such that finite size effects are negligible and a cutoff radius of $4\sigma$ is employed. All of the solid-liquid coexistence simulations are carried out under the $NVT$ ensemble. To produce the $N\sigma T$ ensemble for the crystal, $NPT$ integrators with anisotropic cell fluctuations are used to allow the cell lengths and angle to fluctuate independently. All simulations are carried out over $10^5$ timesteps, 20% of which are used for equilibration with a time step of 1 fs.

The SLE boundaries were obtained by applying the freeze method and the

GDI technique, as described in the Methodology section. Several coexistence points are found and that of ($T_0^*$=1.081, $P_0^*$=5.472) is chosen as the initial known coexistence point required to perform the GDI method, to complete the SLE boundaries. Comparison between the simulation results and the available data from Mastny and de Pablo [52] are presented in Table 2. The simulation results are in good agreement with the SLE boundaries determined by Mastny and de Pablo [52] and are presented in Figure 5. Simulations details and results are presented in Tables 18, 19, 20, 22 and 21 in the Appendix.

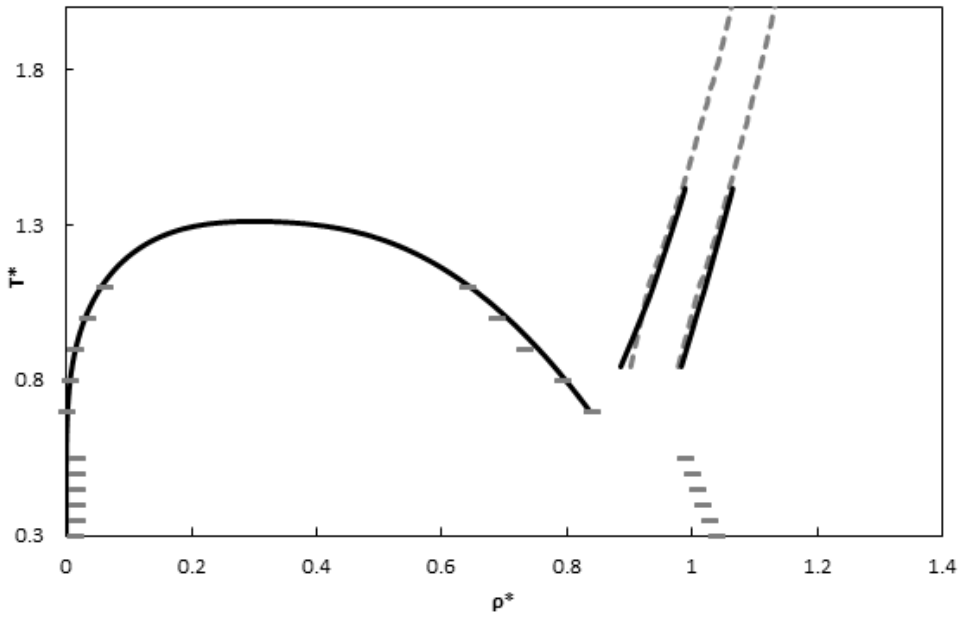


Figure 5: Global phase diagram for Lennard Jones particles showing simulation results (grey symbols) and available correlations (black lines): Mastny and de Pablo [52] for the SLE curves, and Johnson et al [34] for the VLE curve. Simulations results for the VLE and SVE boundaries are from the work of Ramrattan [61]. Simulation results for the SLE are from this work.

Table 2: Solid-liquid coexistence simulation results for the Lennard-Jones system.

| | $T^*$ | $P^*$ | $\varrho^*_{solid}$ | $\varrho^*_{liquid}$ |
|---|---|---|---|---|
| This work | 1.081 | 5.472 | 1.01759 | 0.93969 |
| Work by Mastny and de Pablo [52] | 1.081 | 5.0406 | 1.0185 | 0.93402 |

## 4.2 Benzene

### 4.2.1 Crystal Structure Prediction

Benzene has up to seven polymorphs. We search for the experimental crystal structures from the CCDC and find 19 structures (BENZEN 00-18). From this structural set, we excluded four entries without 3-D coordinates (BENZEN 05, 08, 09, 10). The remaining crystal structures can be clustered. The clustering clearly reveals three clusters of structures: Form II (16, 17), Form III (03, 04) and Form I (0002, 0607, 1115, 1819). In our study we focused on polymorphs I (Form I) and III (Form II or III), which have been unambiguously resolved. Form II and Form III although have different entries in the CCDC, are the same structure observed in different conditions [60].
We perform a global search for possible structures using *CrystalPredictor* [36]. For the repulsion/dispersion interactions representation, the empirical Buckingham potential with the FIT transferable parameters is used. Electrostatic interactions are represented with point charges obtained from QM using the HF/6-31(d,p) level of theory. During the CSP search both of benzene's polymorphs are successfully identified as lattice energy minima. The resulting lattice energy landscape which has 13739 unique structures within +29.86 kJmol$^{-1}$ of the global minimum, and 3019 unique structures within +11.04 kJmol$^{-1}$, as shown in Figure 6. The global minimum of the landscape corresponds to Form III, whilst Form I is ranked 5th, with an energy +1.024 kJmol$^{-1}$ above the global minimum. This is in contrast to experimental evidence that Form I is more stable than Form III at 0K [60]. The agreement between experimental and computational structures is quantified via the root mean squared deviation of the 15-molecule coordination sphere ($rms_{15}$), calculated using Mercury [51].

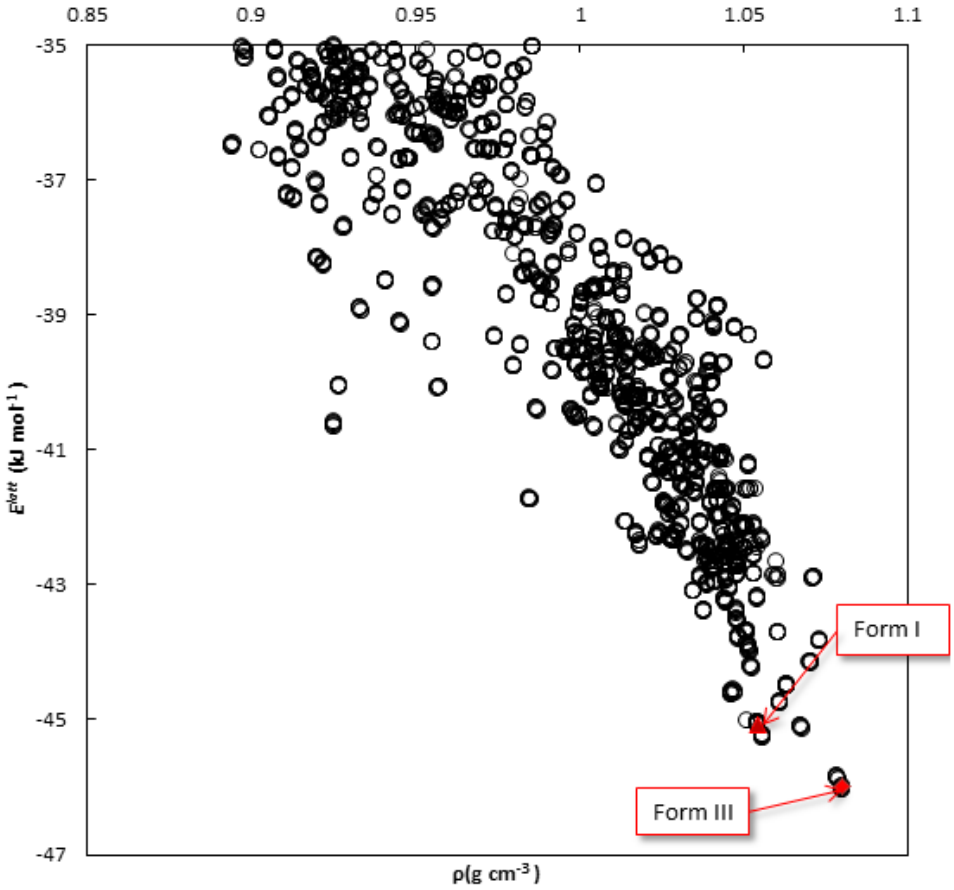


Figure 6: Lattice energy $E^{latt}$(kJmol$^{-1}$) vs density $\varrho$(gcm$^{-3}$) of structures generated during the global search with *CrystalPredictor*. The predicted structures that correspond to known polymorphs of benzene are shown with red.

Form III is is found to have $rms_{15}$=0.106Å, while Form I is found to have an $rms_{15}$ 0.010Å. These differences are visualized in Figure 7 where the experimental and predicted structures are overlaid and additional information for the predicted structures and comparison with the available experimental data is presented in Tables 3&4.

Table 3: Information about the two predicted structures during the CSP global search.

| Structure | $rms_{15}$ | $E^{latt}$(kJmol$^{-1}$) | Rank |
|---|---|---|---|
| Predicted Form I | 0.106 | -46.04 | 1 |
| Predicted Form III | 0.010 | -45.02 | 5 |

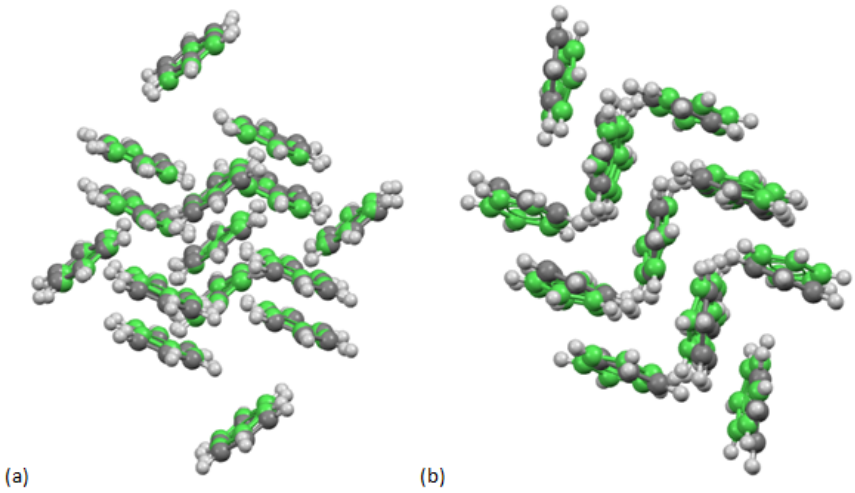


Figure 7: Overlay between predicted (green) and experimental (grey) structures for Form III (a) and Form I (b).

Table 4: Basic information about the two structures corresponding to the experimentally known polymorphs, generated during the global search.

| Structure | density (gcm$^{-3}$) | Space Group |
|---|---|---|
| Experimental Form I | 1.05 | $Pbca$ |
| Predicted Form I | 1.0509 | $Pbca$ |
| Experimental Form III | 1.0793 | $P2_1/c$ |
| Predicted Form III | 1.0796 | $P2_1/c$ |

The global search can be characterised as successful, since both experimental forms of benzene were identified within the lowest-energy and most stable structures.

### 4.2.2 Force fields and methods for molecular simulations

The freeze method is applied for two different force fields, the force field used in *CrystalPredictor* (CSP force field) and the OPLS force field [35] [12].

- CSP Force Field

Initially, the empirical force field used in *CrystalPredictor* is employed.

$$U_{ij}(r) = A_{ij}exp\Big( - \frac{r_{ij}}{B_{ij}}\Big) - \frac{C_{ij}}{r_{ij}^6} + \frac{q_i q_j}{r_{ij}} \tag{10}$$

where $A_{ij}$, $B_{ij}$ and $C_{ij}$ are transferable parameters which describe same site/element interactions and are fitted to experimental data
$q$ is the charge variable
$r_{ij}$ is the interatomic distances between two molecules.

The repulsion/dispersion interactions are represented using the semi-empirical Buckingham potential with transferable parameters available in literature (FIT parameters). Electrostatic interactions are represented with Coulombic contribution with point charges obtained from QM (HF/6-31(d,p) level of theory).

- OPLS Force Field

As an alternative to the CSP force field, the OPLS force field [35] is considered with parameters listed by Cacelli et al [11][12] who employed a model intermolecular potential derived from quantum mechanical calculations. The *ab initio* database includes approximately 200 geometries of the benzene dimer with interaction energies computed at the MP2 level of theory. The OPLS force field was parameterized for benzene using *ab initio* interaction energies and was tested along with the optimized potentials for liquid simulation (OPLS) benzene force field, which was originally parameterized using heat of vaporization and liquid density. For this model, a rigid version was used for solid benzene and compared densities and lattice parameters were compared with experimental data resulting in good agreement. The atom-atom potential is given by:

$$U_{ij}(r) = 4\epsilon_{ij}\Big[\Big(-\frac{\sigma_{ij}}{r_{ij}}\Big)^{12} - \Big(-\frac{\sigma_{ij}}{r_{ij}}\Big)^{6}\Big] + \frac{q_i q_j}{r_{ij}} \tag{11}$$

where $\sigma$ and $\epsilon$ are adjustable parameters
$q$ is the charge variable
$r_{ij}$ is the interatomic distances between two molecules
The repulsion/dispersion interactions are represented by the Lennard-Jones potential with adjustable parameters $\sigma$ and $\epsilon$. For the Coulombic contribution one charge variable is used. There are five adjustable parameters in total $\sigma$ and $\epsilon$ for carbon and hydrogen, and one charge variable, as shown in Table 5.

Table 5: Adjustable parameters $\sigma$ and $\epsilon$ for carbon and hydrogen [35] [11].

| OPLS Intermolecular Parameters used in Rigid Models | |
|---|---|
| $\sigma_C$ (Å) | 3.55 |
| $\epsilon_C$(kJmol$^{-1}$) | 0.293 |
| $q_C$ (e) | -0.115 |
| $\sigma_H$ (Å) | 2.47 |
| $\epsilon_H$(kJmol$^{-1}$) | 0.125 |
| $q_H$ (e) | 0.115 |

- Simulation details

For each identified polymorph the melting point calculation is carried out starting from the computationally obtained structures. Initial crystal configurations are based on a $6x6x6$ super cell for the global minimum structure and a $5x4x3$ super cell for the 5th structure built from the predicted unit cells at 0K and 0 Pa, with a total of 432 molecules. The liquid simulations also contain 432 molecules.
Throughout this work, all simulations are performed using the well-known DL_POLY software for MD simulations. For determining the hysteresis region and calculating single phase properties, simulations are conducted under the $N\sigma T$ and $NPT$ ensembles. All the solid-liquid coexistence simulations are carried out under the $NVT$ ensemble. To produce the $N\sigma T$ ensemble for the crystal, $NPT$ integrators with anisotropic cell fluctuations are used

to allow the cell lengths and angles to fluctuate independently. Isotropic fluctuations are used for liquid simulations with the Nosé-Hoover thermostat and barostat. For all simulations, the repulsion/dispersion interactions are truncated at 15.0 Å and standard long range corrections are used to correct energy and pressure. Long-range Coulombic interactions are calculated with the Ewald summation scheme. For the pure-phase simulations, a time step of 1 fs is used throughout the runs. For the solid-liquid coexistence simulations, a time step of 0.5 fs is used.
Heating of the solid and cooling of the liquid with a temperature increment of 20K is used to determine the hysteresis region. We use incremental heating/cooling in the melting and crystallization transition simulations, that is, the starting configuration of each simulation is the last configuration of the previous simulation at a lower or higher temperature. The simulation temperatures start at 0K and step up to 360K in increments of 20K. Trial an error based on removal of molecules as explained in the Methodology section is used to determine a solid-liquid coexistence point, necessary for the application of our melting point prediction method.

The methodology described above is applied for three systems and two force fields are employed. The systems under investigation for melting point prediction are the global minimum structure with the CSP force field, the global minimum structure and OPLS force field and the 5th structure and OPLS force field.

Based on the simulation techniques used here, it is not possible to asses the relative stability of the two computed structures. As a result, the calculated data obtained with both structures are compared to experimental data from the most stable form, Form I.

### 4.2.3 Melting point prediction for the global minimum structure and CSP force field

1. The method is initially carried out for the global minimum crystal structure which corresponds to the experimental benzene Form III and the force field used in $CrystalPredicto$r is employed.
2. Initial crystal configurations are based on a $6x6x6$ super cell built from the predicted unit cell of the global minimum structure at 0K and 0 Pa, with a total of 432 molecules. The liquid simulation box also contains 432

molecules.
3. Heating of the solid and cooling of the liquid with a temperature increment of $\Delta T$=20K determine a hysteresis region (Figure 8). The simulations temperatures start at 0K and step up to 310K in increments of 20K. Trial an error based on removal of molecules determine a solid-liquid coexistence point, necessary for the application of our melting point prediction method.

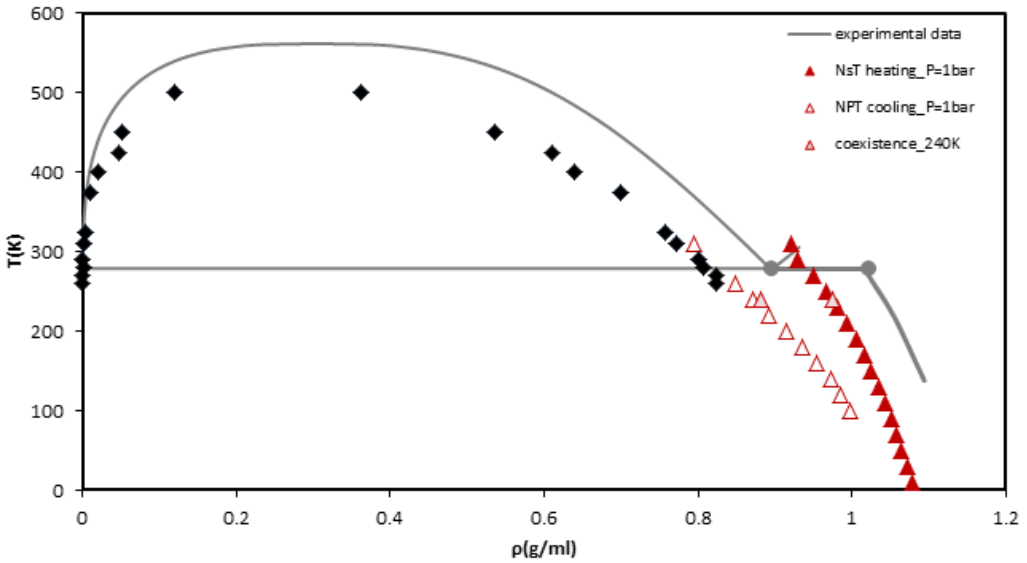


Figure 8: Diagram for the global minimum structure and CSP force field showing the heating (solid red triangles) and cooling (open red triangles) isobars at P=1 bar. Experimental data [74] [20] [13] [62] are represented with grey lines and grey points. Simulations carried out to determine the VLE region are represented with black points.

4. For the construction of the $SLE$ box, the solid configuration at $T$=290K is replicated four times towards the z-direction. After densities are fixed, the resulting SLE configuration contains 1714 molecules, in the solid region $N_{solid} = 864$ molecules and in the liquid region $N_{liquid} = 850$ molecules.

For the solid-liquid coexistence simulations, we choose the temperature of 240 K which is within the hysteresis region determined in the previous step. Initially, the system is relaxed under the $NVT$ ensemble with molecules in the solid region remaining frozen. The unfrozen half melts into a liquid while

the frozen particles remain stationary. Finally, the stationary molecules are released and the entire system is allowed to equilibrate under the canonical $NVT$ ensemble for a total of $1.7 * 10^6$ time steps. The $NVT$ run reaches 1.5 ns in real time. From this 0.2 ns is reached with a time step of 0.0005 ps (for $4 * 10^5$ steps ) and the remaining 1.3 ns with a time step of 0.001ps (for $1.3 * 10^5$ steps).

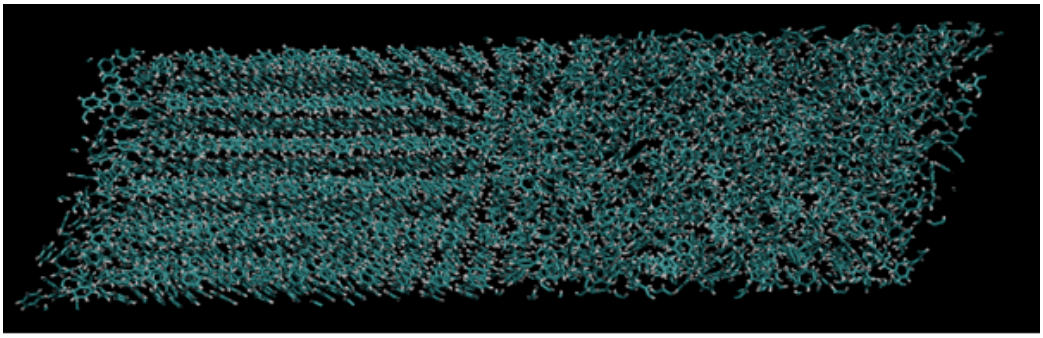

Figure 9: Solid-liquid coexistence configuration for the global minimum structure and CSP force field at T=240K after 1.5ns under the $NVT$ ensemble.

5. From the $NVT$ simulation the density profile is obtained by calculating the number of molecules that have their centre of mass within a certain bin:

$$\varrho(z) = \frac{N(z)}{V(z)} \tag{12}$$

where:
$\varrho(z)$ is the density within the bin that contains position z,
$N(z)$ is the number of molecules with centre of mass within this bin, and
$V(z)$ is the volume of this bin.

The density profile showed in Figure 10 is obtained using a bin width of 1.5 Åand it shows two well defined plateaus corresponding to the coexisting

phases.

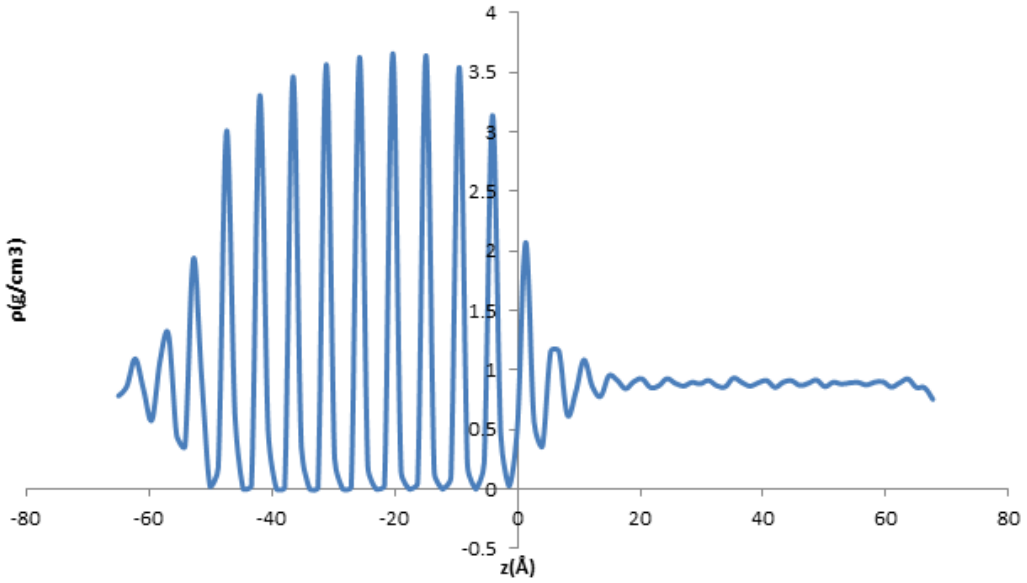


Figure 10: Density profile of the SLE simulation box for the global minimum structure and CSP force field

The Radial Distribution Function (RDF) of each phase is obtained confirming that solid-liquid coexistence is maintained at these conditions. The RDF for the solid is based on the molecules positioned between $z = -40$Å and $z = -10$Å in the simulation box. That of the liquid phase is based on the range $z = 20$Å and $z = 60$Å.

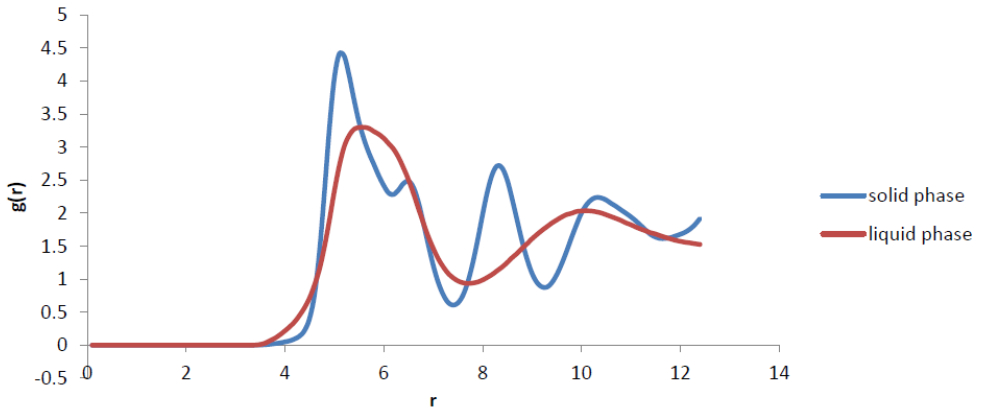


Figure 11: Radial Distribution Functions g(r) for the two phases of the co-existence system for the global minimum structure and CSP force field. The blue line corresponds to the RDF of the solid phase and the red line to the one of the liquid phase.

As mentioned above, the density profile and RDFs confirm the presence of two phases. The density for the solid region between $z = -40$Å and $z = -10$Å is found to be 0.9618 g cm$^{-3}$ and the density for the liquid phase between $z = 20$Å and $z = 60$Å is found to be 0.8824 g cm$^{-3}$.

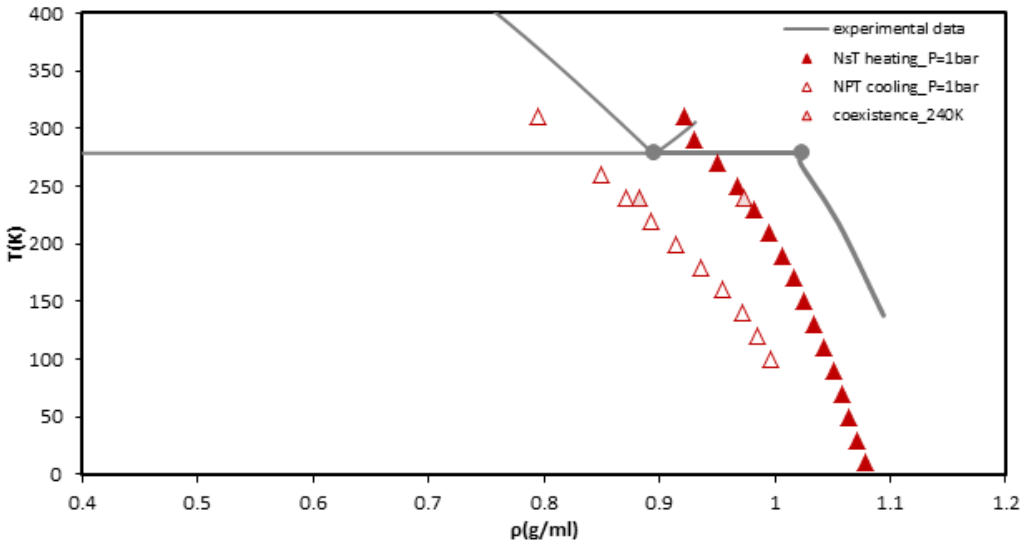


Figure 12: Solid-liquid coexistence point at $T$=240K (light red triangles) along the heating expansion (solid red triangles) and cooling compression(open red triangles), isobars at $P$=1 bar. Experimental data are represented with grey lines and grey points.

The output pressure of the coexistence $NVT$ simulation at $T$=240K obtained as the normal component of the pressure tensor is found to be $P$=0.290275 katm. The pressure obtained from the single phase $NVT$ simulation of the solid is found to be $P$=0.78795 katm and for the liquid $P$=1.1228 katm. These results are summarised in Table 11.

Table 6: Output pressure results from the coexistence $NVT$ simulation and the single solid and liquid phase $NVT$ simulations.

| $T$(K) | $P_{coexistence}$(katm) | $P_{solid\ single\ phase}$(katm) | $P_{liquid\ single\ phase}$(katm) |
|---|---|---|---|
| 240 | 0.290275 | 0.78795 | 1.1228 |

This kind of variation in pressure is expected because $P$ is a very steep function of density. Calculated pressures in systems with an interface, such as this, can be expected to have large errors. Pressure is calculated using the average pairwise forces and even a small shift in the molecules can have

a massive effect on the forces.This aspect warrants further investigation but this is beyond the scope of this study.
As mentioned previously in the methodology section, locating a coexistence point is necessary for applying the Gibbs-Duhem integration and generate the complete phase diagram of the substance of interest. The conditions reported in Table 6 ensure maintenance of the solid-liquid coexistence and the RDFs show that solid and liquid phase simulations are consistent with the two-phase simulations. Therefore it is this ($T$,$P$) point and the corresponding single phase configurations that are used as the initial point for the Gibbs-Duhem integration.

The corresponding single phase configurations must be obtained from the coexistence system. For this reason a method is developed to allow us to extract unit cells from solid simulations, which is very useful for the application of our method as well as investigating potential phase transitions. The method provides a unit cell from the solid region of the coexistence system. Replication of this unit cell to a simulation box ensures that the solid configuration necessary for the independent single phase simulations has the exact same density and structure as the solid of the coexistence.

For the liquid phase, a similar method is developed which allows the extraction of clusters from the liquid region of the coexistence system. The centre of mass for each molecule within the liquid region is located. This results a configuration with the exact density of the liquid region of the coexistence system which is necessary for the independent single phase simulations.

Independent single phase $N\sigma T$ and $NPT$ simulations for the solid and the liquid phase are carried out over $10^5$ time steps, 20% of which are used for equilibration, as described in the methodology section. For the integration a temperature increment of $\Delta T$=3 K is applied. Results of the integration are shown in Table 7. Additionally, in Figures 13 and 14, the ($T$,$\varrho$) phase diagram is shown compared to the available experimental data.

Table 7: Single phase $N\sigma T$ and $NPT$ simulation results for the solid and the liquid phase respectively for the system of the global minimum structure and the CSP force field.

| $T$(K) | $\varrho_{solid}$(g/cm$^3$) | $\varrho_{liquid}$(g/cm$^3$) |
|---|---|---|
| 240 | 0.97393 | 0.88245 |
| 243 | 0.975845 | 0.88297 |
| 246 | 0.97712 | 0.883675 |
| 249 | 0.978876 | 0.885431 |
| 252 | 0.97586 | 0.883956 |
| 255 | 0.977235 | 0.885723 |
| 258 | 0.979002 | 0.887279 |
| 261 | 0.980882 | 0.889046 |
| 264 | 0.982752 | 0.890766 |
| 267 | 0.984572 | 0.892136 |
| 270 | 0.986328 | 0.893956 |

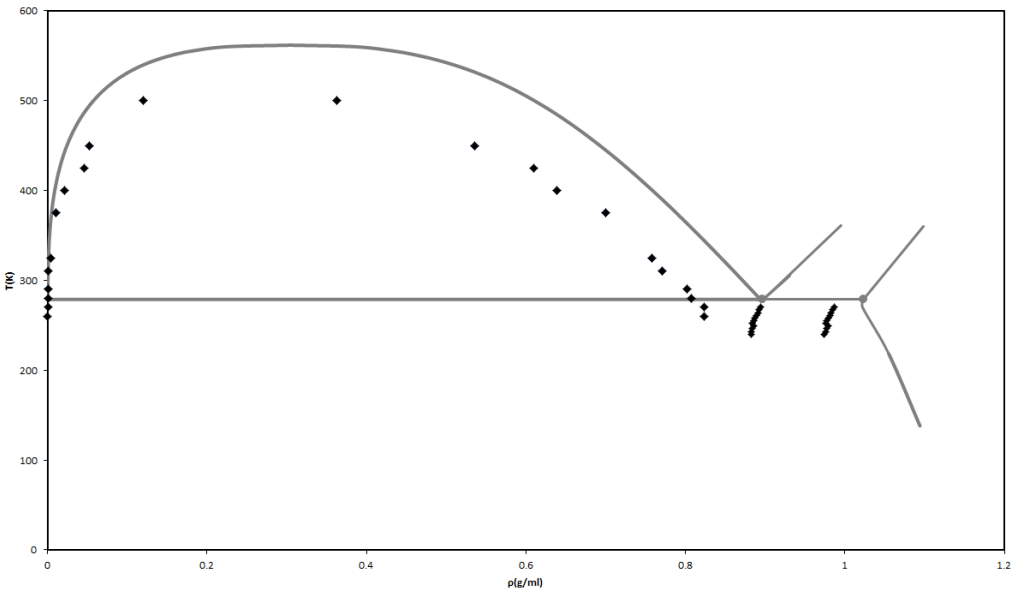


Figure 13: Phase diagram for the global minimum structure and the CSP force field showing the boundaries of phase transition. Simulation data for the VLE and SLE are represented with black diamonds. Experimental data are represented with grey lines and points [74] [20] [13].

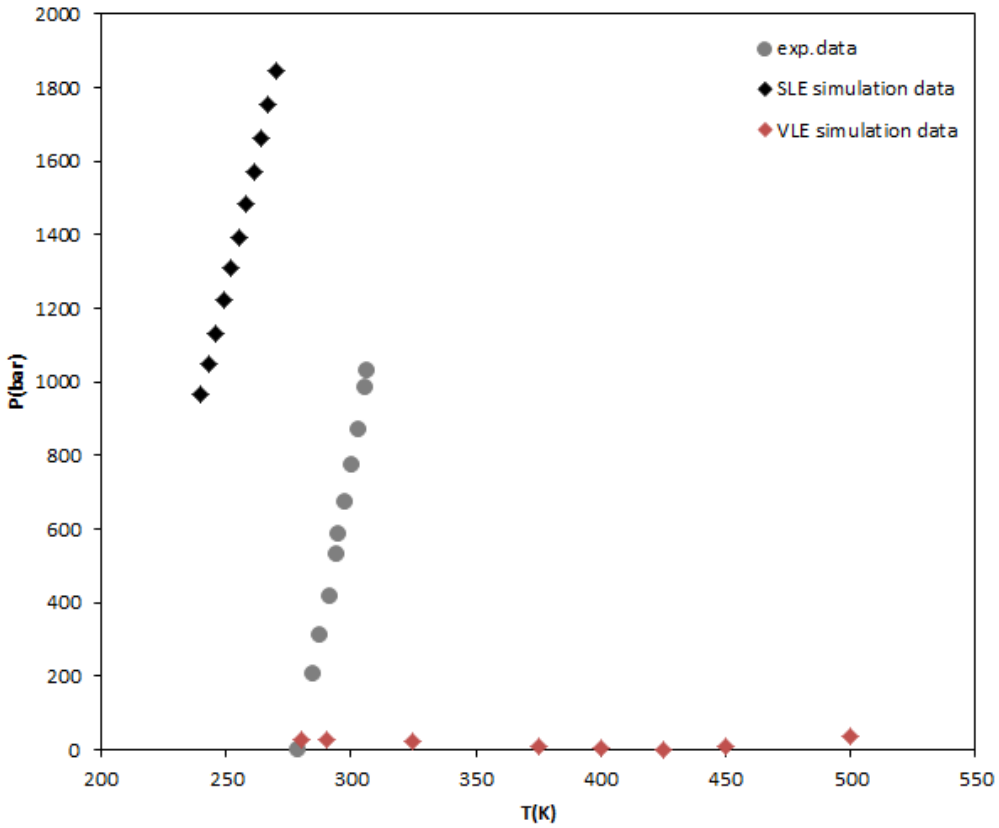


Figure 14: $P$-$T$ melting line for the global minimum structure and the CSP force field. Simulation data for the SLE are represented with black diamonds, simulations data for the VLE are represented with red diamonds and experimental data are represented with grey circles [74][3].

As the $(T,\varrho)$ phase diagram of the global minimum structure and the CSP force field shows there is a correct trend for the phase transition boundaries. There is clearly an underestimation of the triple point of at least 50 K. The differences between solid and liquid densities are also underestimated.
The ($P$-$T$) diagram shows an underestimation of temperature for given pressure compared to the available data. This systems results in a melting point of approximately $T_m$=210 K at $P$=1 bar when the experimental value is $T_{m\,\mathrm{exp}}$=278.68 K at $P$=1 bar.
Accuracy of the prediction and agreement with the experimental data depends on the force field used. For this reason a different force field is employed

in order to investigate the possibility of better predictions.

### 4.2.4 Melting point prediction for the global minimum structure and OPLS force field

1. The method is initially carried out for the global minimum crystal structure which corresponds to the experimental benzene Form III and and the rigid version of the OPLS force field is employed.
2. Initial crystal configurations are based on a $6x6x6$ super cell built from the predicted unit cell of the global minimum structure at 0K and 0 Pa, with a total of 432 molecules. The liquid simulation box also contains 432 molecules.
3. Initially, the solid is heated at $T$=2K and then at $T$=10K. Temperature is gradually risen at $T$=138K, $T$=218K, $T$=270K, $T$=290K and $T$=300K. After that a temperature increment of $\Delta T$=20K is used up until $T$=360K where a sudden decrease in density is observed, indicating formation of a liquid. These temperatures are chosen to allow comparison with the available simulation and experimental data for benzene and the rigid version of the OPLS force field [35] [11][20]. Comparison is shown in Tables 8, 9 and 10 and although there is a slight overestimation of density noted, the agreement between the available experimental and simulation data is acceptable. Cooling of the liquid follows at $T$=312K, $T$=300K, $T$=280K, $T$=270K, $T$=240K, $T$=210K and $T$=180K. The resulting heating and cooling isobars determine the hysteresis region, as shown in Figure 15.

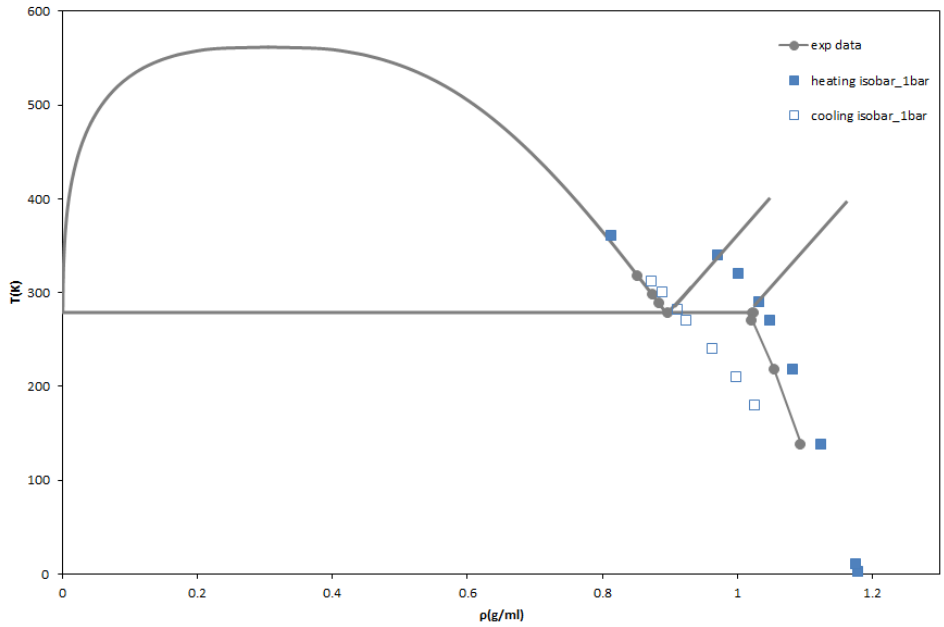


Figure 15: Diagram for the global minimum structure and OPLS force field showing the heating (solid blue boxes) and cooling (blue open boxes) isobars at $P$=1 bar. Experimental data are represented with grey lines and points [62] [74] [20] [13].

Table 8: Benzene Crystal Data at $T$=218 K and $P$=1 bar.

| | experiment | Eike and Maginn [22] | Cacelli et al [11] | this work |
|---|---|---|---|---|
| $\varrho$(g/cm$^3$) | 1.055 | 1.039 | 1.060 | 1.083 |

4. For the construction of the SLE box, the solid configuration at T=290K is replicated four times towards the z-direction. After densities are fixed by removal of molecules, the resulting SLE configuration contained 1673 molecules, in the solid region $N_{solid} = 864$ molecules and in the liquid region $N_{liquid} = 809$ molecules.
For the solid-liquid coexistence simulations, we choose the temperature of 240K which is within the hysteresis region determined in the previous step. The resulting system is relaxed under the $NVT$ ensemble with molecules in the solid region remain stationary while the unfrozen half melts into a

Table 9: Benzene Crystal Data comparison between experimental data and simulation data by Cacelli et al and this work.

| | Experimental Data | Cacelli et al [11] | this work |
|---|---|---|---|
| $T$(K) | | $\varrho(g/cm^3)$ | |
| 138 | 1.094 | 1.086 | 1.124 |
| 218 | 1.055 | 1.060 | 1.083 |
| 270 | 1.022 | 1.007 | 1.048 |

Table 10: Comparison between simulation data by Cacelli et al and this work for the liquid phase.

| | Cacelli et al [11] | this work |
|---|---|---|
| $T$(K) | $\varrho$(g/cm$^3$) | |
| 281 | 0.905 | 0.912 |
| 312 | 0.857 | 0.873 |

liquid. After that, the stationary molecules are released and the entire system is allowed to equilibrate under the canonical $NVT$ ensemble for a total of $1.7 * 10^6$ time steps. The $NVT$ run reached 1.5 ns in real time. From this 0.2 ns was reached with a time step of 0.0005 ps (for $4 * 10^5$ steps) and the rest 1.3 ns with a timestep of 0.001ps (for $1.3 * 10^5$ steps).

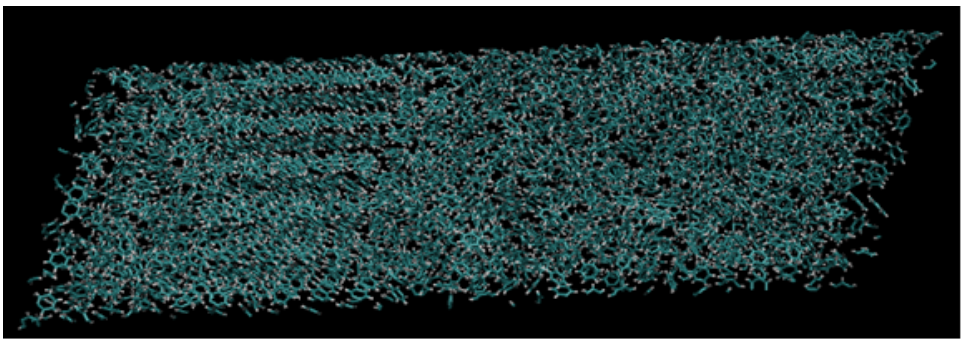

Figure 16: Solid-liquid coexistence configuration at $T$=250K after 1.5ns under the $NVT$ ensemble for the global minimum structure and the OPLS force field.

5. From the $NVT$ simulation the density profile, as mentioned in the previous section, is obtained by calculating the number of molecules that have their centre of mass within a certain bin. For the density profile showed in Figure10 a bin width of 1.5Å is used. Two well defined plateaus are observed, indicating the coexistence of two phases.

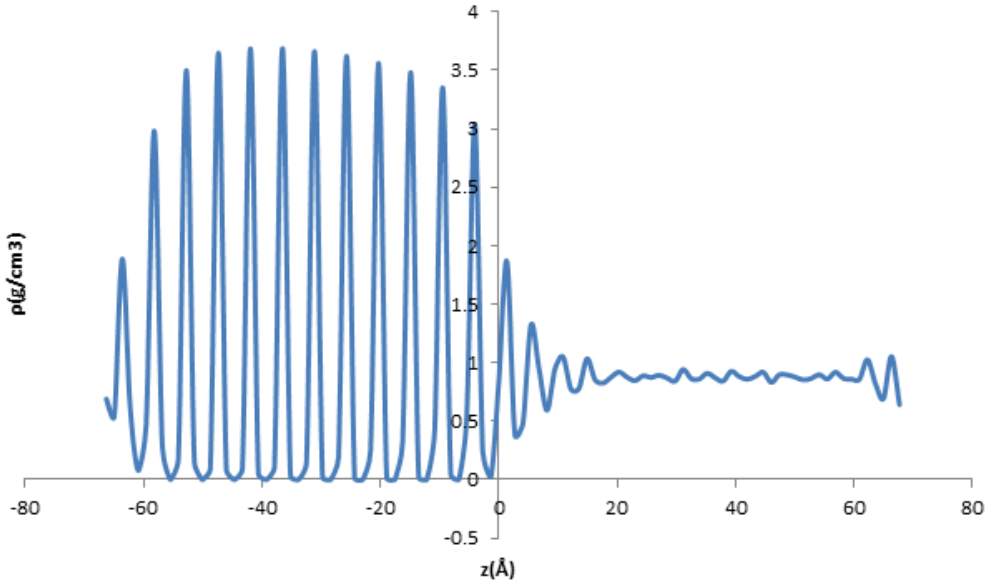


Figure 17: Density profile of the $SLE$ simulation box

The RDF of each phase is obtained confirming that solid-liquid coexistence is maintained at these conditions. The RDF for the solid is based on the molecules positioned between $z = -40$Å and $z = -10$Å in the simulation box. That of the liquid phase is based on the range $z = 20$Å and $z = 60$Å.

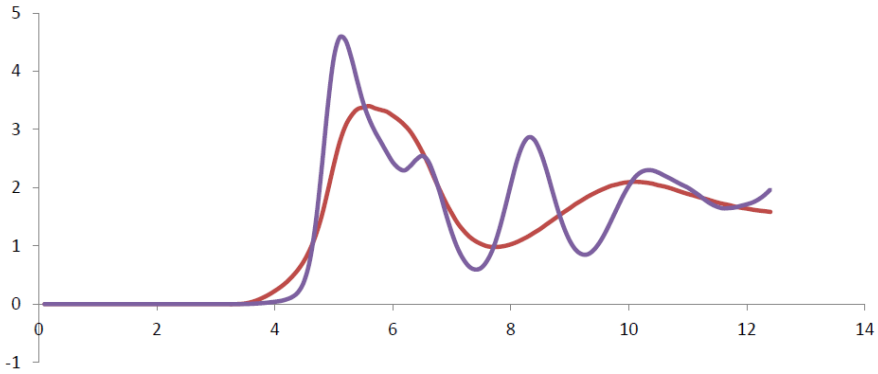


Figure 18: Radial Distribution Function g(r) for the two phases of the coexistence system. The purple line corresponds to the RDF of the solid phase and the red line to the one of the liquid phase.

The density profile and RDFs confirm the presence of two phases. The density for the solid region between $z = -50$Å and $z = -20$Å is found to be 0.9589 g cm$^{-3}$ and the density for the liquid phase between $z = 30$Å and $z = 50$Å is found to be 0.8789 g cm$^{-3}$. The RDF of the two phases is obtained conforming that solid-liquid coexistence is maintained at these conditions.

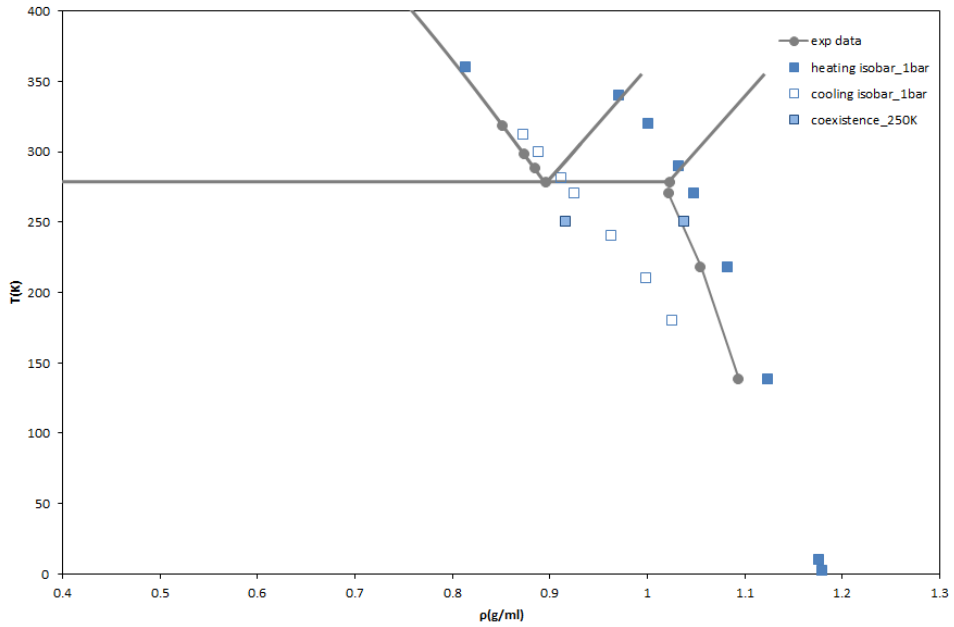


Figure 19: Solid-liquid coexistence point at $T$=250K (light blue boxes) along the heating (solid blue boxes) and cooling (open blue boxes) expansion and compression respectively, isobars at $P$=1 bar. Experimental data are represented with grey lines and points [74] [20] [13].

The output pressure of the coexistence $NVT$ simulation at $T$=240K obtained as the normal component of the pressure tensor is found to be $P$=106.38 katm. The pressure obtained from the single phase $NVT$ simulation of the solid is found to be $P$=1.8775 katm and for the liquid $P$=1.0152 katm. These results are reported in Table 11.

Table 11: Output Pressure results from the coexistence $NVT$ simulation and the single solid and liquid phase $NVT$ simulations.

| $T$(K) | $P_{coexistence}$(katm) | $P_{solid\ single\ phase}$(katm) | $P_{liquid\ single\ phase}$(katm) |
|---|---|---|---|
| 240 | 106.38 | 1.8775 | 1.0152 |

There is clearly a large difference in pressures of the coexistence system compared to the single phase simulations. This variation can be expected in such systems. Additionally, $P$ is a very steep function of density and is

calculated using the average pairwise forces. As a result, even a small shift in the molecules can have a massive effect on the forces.
The conditions reported in Table 11 ensure maintenance of solid-liquid coexistence. Therefore it is this point and the corresponding single phase configurations that are used for the initial point for the Gibbs-Duhem integration. Solid and liquid configurations for the independent single phase simulations are extracted from the coexistence box according to the methods described in paragraph 5 for the global minimum structure and CSP force field.
Simultaneous single phase $N\sigma T$ and $NPT$ simulations for the solid and the liquid phase are carried out over $10^5$ timesteps, 20% of which are used for equilibration, as described in the methodology section. For the integration a temperature increment of $\Delta T$=3K is applied. Results of the integration are shown in Table 12. Additionally, in Figures 20 and 21, the ($T$,$\varrho$) and ($P$,$T$) phase diagrams are shown compared to the available experimental data.

Table 12: Single phase simulation results for the global minimum structure and the OPLS force field.

| $T$(K) | $\varrho_{solid}$(g/cm$^3$) | $\varrho_{liquid}$(g/cm$^3$) |
|---|---|---|
| 250 | 1.0374 | 0.9161 |
| 253 | 1.039315 | 0.91747 |
| 256 | 1.04059 | 0.91922 |
| 259 | 1.042346 | 0.92086 |
| 262 | 1.044216 | 0.92258 |
| 265 | 1.045796 | 0.92416 |
| 268 | 1.047466 | 0.92603 |
| 271 | 1.048916 | 0.92785 |
| 274 | 1.05018 | 0.92914 |
| 277 | 1.0518 | 0.93081 |
| 280 | 1.05334 | 0.93268 |

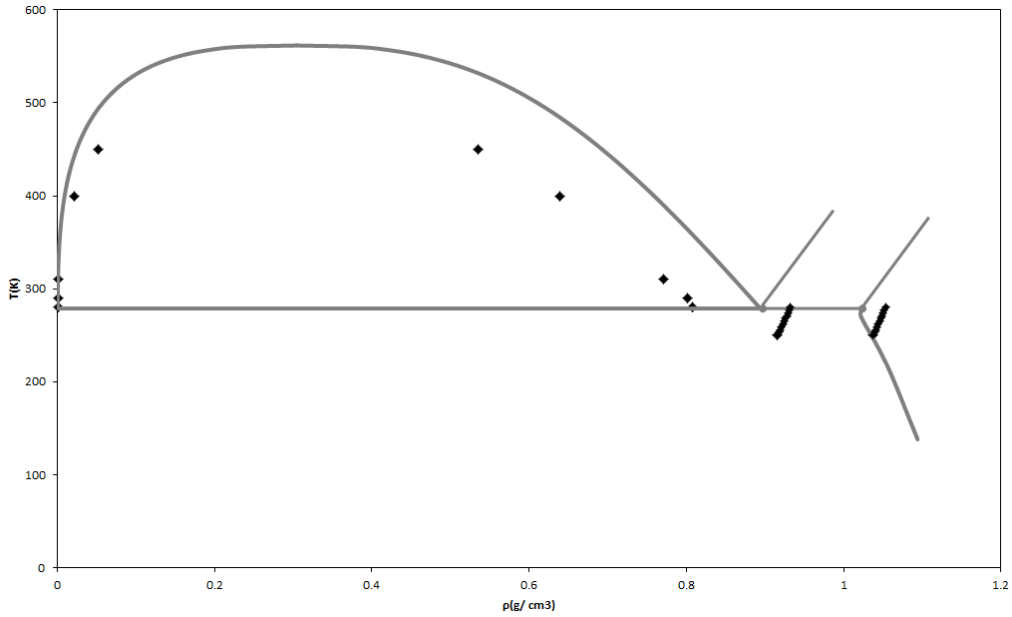


Figure 20: Phase diagram for the global minimum structure of benzene using the OPLS force field. Simulation results for the VLE and SLE are represented with black diamonds. Experimental data are represented with grey lines [74] [20] [13].

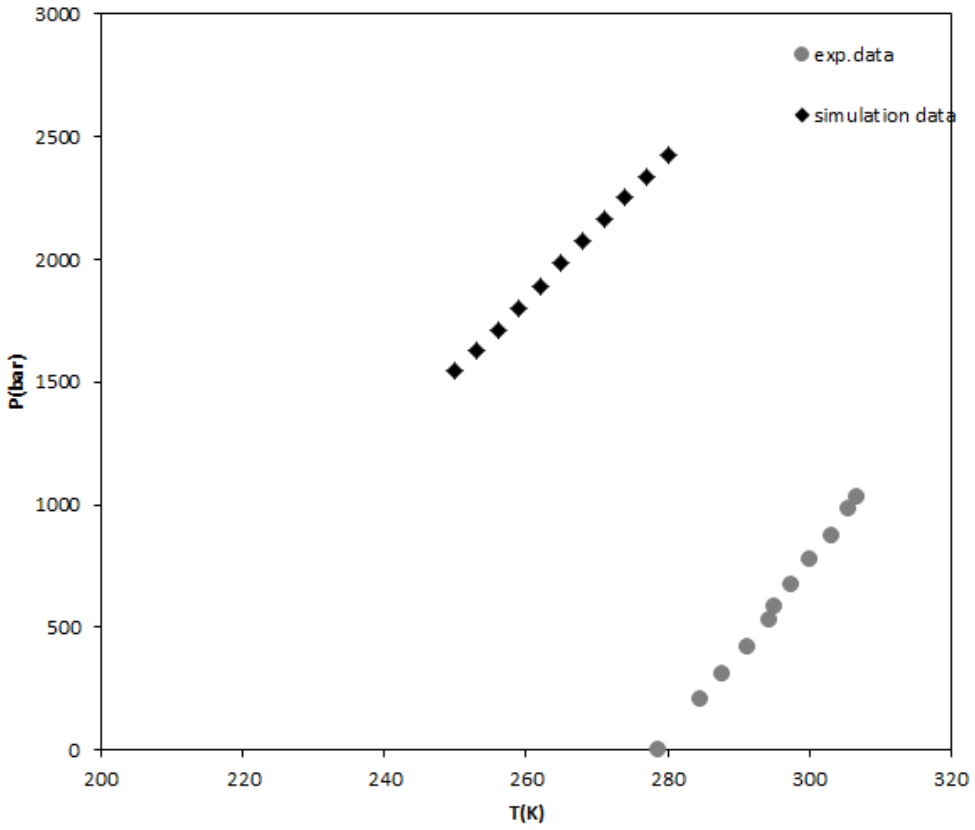


Figure 21: $P$-$T$ melting line for the global minimum structure and the OPLS force field. Simulation data for the SLE are represented with black diamonds and for the VLE with red diamonds. Experimental data are represented with grey points [74][3].

The ($T$,$\varrho$) phase diagram of the global minimum structure and the OPLS force field, appears to be shifted below by approximately 40-50 K. Although there is a correct trend for the phase transition boundaries, there is clearly an underestimation of the triple point of at least 40-50 K. Difference between solid and liquid densities appear to be larger and better compared to the global minimum structure and the CSP force field phase diagram. There is also a slightly worse performance on the VLE region.
The ($P$-$T$) diagram shows an underestimation of temperature for given pressure compared to the available data. This systems results a melting point of approximately $T_m$=200K at $P$=1bar when the experimental value is $T_{m\,\mathrm{exp}}$=278.68K

at $P$=1bar.
As mentioned above, the accuracy of the prediction is highly depended on the force field employed. It also depends on the structure used for the calculation. For this reason, the 5th generated structure is going to be investigated by employing the OPLS.

### 4.2.5 Melting point prediction for the 5th ranked structure and OPLS force field

1. The method is initially carried out for the 5th ranked crystal structure which corresponds to the experimental benzene Form I and the rigid version of the OPLS force field is employed.
2. Initial crystal configurations are based on a $5x4x3$ super cell built from the predicted unit cell of the global minimum structure at 0K and 0 Pa, with a total of 432 molecules. The liquid simulation box also contains 432 molecules.
3. Initially, the solid is heated at $T$=2K and then at $T$=10K. Temperature is gradually risen at $T$=110K, $T$=130K, $T$=138K, $T$=150K, $T$=170K, $T$=190K, $T$=218K, $T$=270K and $T$=300K. As for the system of the global minimum structure and OPLS force field, these temperatures are chosen to allow comparison with the available simulation and experimental data for benzene and the rigid version of the OPLS force field [35] [11][20]. Comparison is shown in Tables 15, 14 and 13 and there is reasonable agreement between the available experimental and simulation data. Also, smaller temperature increments at $T = 110 - 190$K are followed to locate a possible solid-solid transition. While heating at constant pressure $P$=1bar from $T$=2K up until $T$=150K density decreases as expected. At $T$=170K there is a sudden increase of density, indicating a possible solid-solid transition. The solid configuration is heated until $T$=300K where a decrease in density is observed, indicating the formation of a liquid. Gradual cooling of the system now in the liquid phase begins at $T$=312K and continues at $T$=280K, $T$=250K and $T$=220K. The resulting heating and cooling isobars determine the hysteresis region as shown in Figure 22. Removal of molecules determine a solid-liquid coexistence point, necessary for the application of the melting point prediction method.

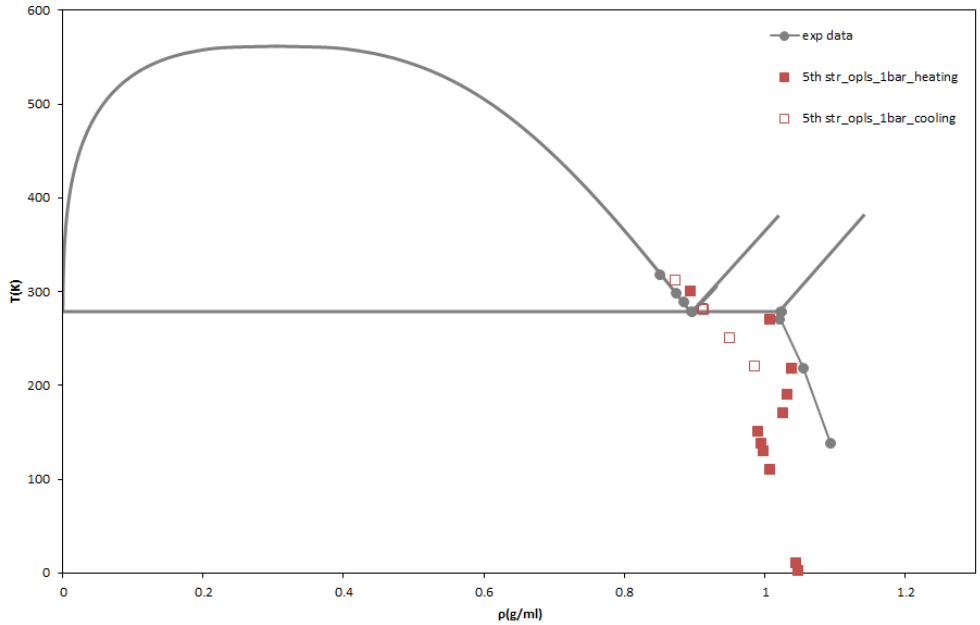


Figure 22: Diagram showing the heating (solid red boxes) and cooling (red open boxes) isobars at $P$=1 bar. Experimental data are represented with grey lines and points [74] [20] [13].

Table 13: Benzene Crystal Data at $T$=218 K and $P$=1 bar.

| | experiment | Eike and Maginn [22] | Cacelli et al [11] | this work |
|---|---|---|---|---|
| $\varrho$(g/cm$^3$) | 1.055 | 1.039 | 1.060 | 1.038 |

Table 14: Benzene Crystal Data comparison between experimental data and simulation data by Cacelli et al and this work.

| | Experimental Data | Cacelli et al [11] | this work |
|---|---|---|---|
| $T$(K) | | $\varrho$(g/cm$^3$) | |
| 138 | 1.094 | 1.086 | 0.996 |
| 218 | 1.055 | 1.060 | 1.038 |
| 270 | 1.022 | 1.007 | 1.007 |

Table 15: Comparison between simulation data by Cacelli et al and this work for the liquid phase.

| | Cacelli et al [11] | this work |
|---|---|---|
| $T$(K) | $\varrho$(g/cm$^3$) | |
| 281 | 0.905 | 0.912 |
| 312 | 0.857 | 0.873 |

4. For the next step, the construction of the $SLE$ box is necessary. For this, the solid configuration at T=290K is replicated for times towards the z-direction. After densities are fixed with removal of molecules, the resulting SLE configuration contains 1679 molecules, in the solid region $N_{solid} = 864$ molecules and in the liquid region $N_{liquid} = 815$ molecules.
For the solid-liquid coexistence simulations, we choose the temperature of 240K which is within the hysteresis region determined in the previous step. After the liquid region melts while molecules in the solid region remain frozen, all molecules are released and the system is relaxed under the $NVT$ ensemble for a total of $1.7 * 10^6$ time steps. The $NVT$ run reached 1.5 ns in real time. From this 0.2 ns was reached with a time step of 0.0005 ps (for $4 * 10^5$ steps) and the rest 1.3 ns with a time step of 0.001ps (for $1.3 * 10^5$ steps).

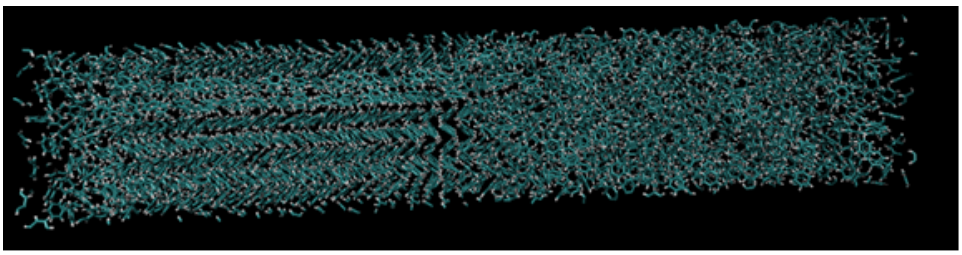

Figure 23: Solid-liquid coexistence configuration at $T$=250K after 1.5ns under the $NVT$ ensemble for the 5th structure and OPLS force field.

5. From the $NVT$ simulation the density profile was obtained by calculating the number of molecules that have their centre of mass within a certain bin. For the density profile shown in Figure 24 a bin width of 2Å.

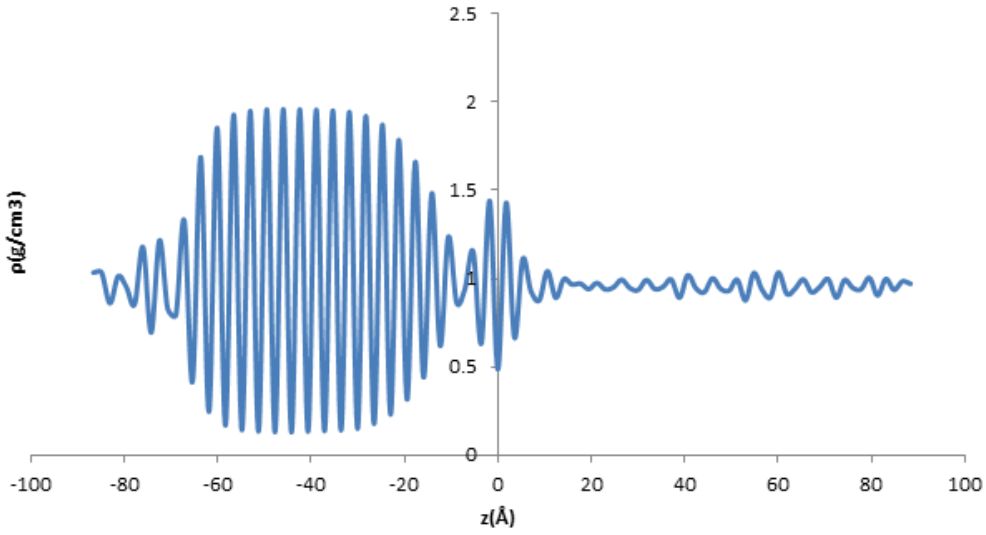


Figure 24: Density profile of the SLE simulation box for the 5th structure and the OPLS force field

The RDFs of the two phases are obtained confirming that solid-liquid coexistence is maintained at these conditions. The RDF for the solid is based on the molecules positioned between $z = -55$Å and $z = -25$Å in the simulation box. That of the liquid phase is based on the range $z = 20$Å and $z = 50$Å.

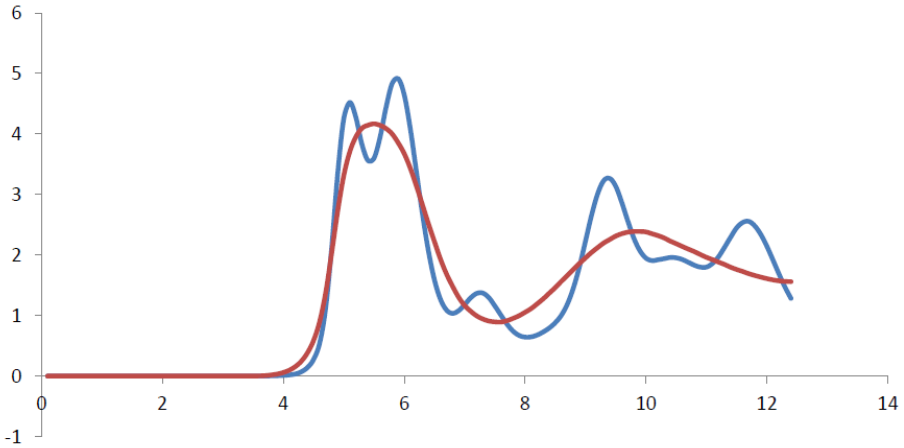


Figure 25: Radial Distribution Function for the two phases of the coexistence system. The blue line corresponds to the RDF of the solid phase and the red line to the one of the liquid phase.

Both density profile and RDF confirm the coexistence of two phases. The density for the solid region between $z = -55$Å and $z = -25$Å is found to be 1.0328 g cm$^{-3}$ and the density for the liquid phase between $z = 30$Å and $z = 50$Å is found to be 0.9555 g cm$^{-3}$. The RDF of the two phases is obtained conforming that solid-liquid coexistence is maintained at these conditions.

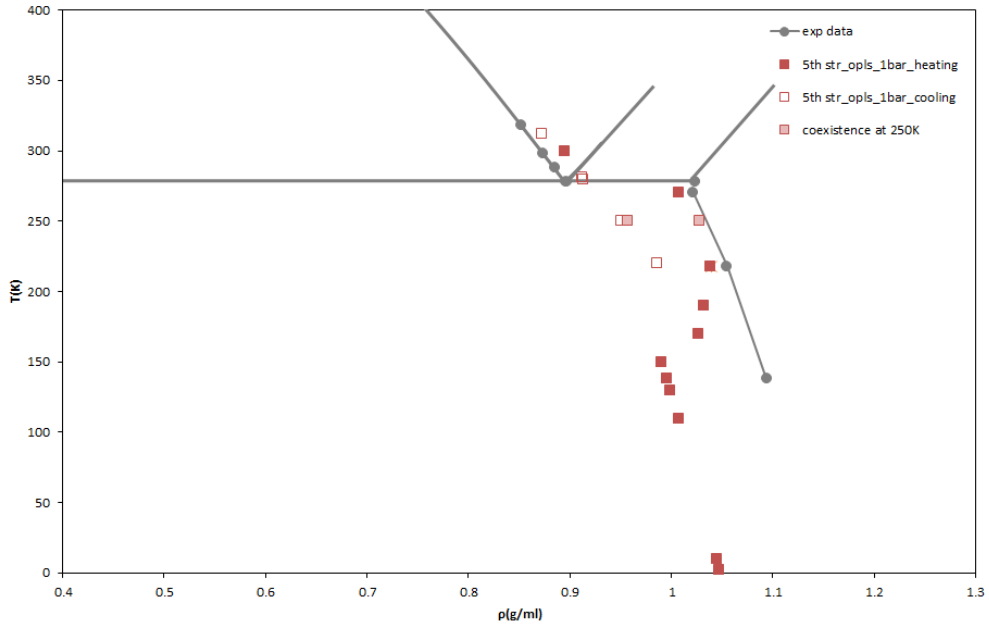


Figure 26: Solid-liquid coexistence point at $T$=250K (light red boxes) along the heating (solid red boxes) and cooling (open red boxes) expansion and compression respectively, isobars at $P$=1 bar. Experimental data are represented with grey lines and points [74] [20] [13].

The output pressure of the coexistence $NVT$ simulation at $T$=240K obtained as the normal component of the pressure tensor was found to be $P$=0.1258 katm. The pressure obtained from the single phase $NVT$ simulation of the solid was found to be $P$=0.070598 katm and for the liquid $P$=0.14087 katm.

Table 16: Output Pressure results from the coexistence $NVT$ simulation and the single solid and liquid phase $NVT$ simulations.

| $T$(K) | $P_{coexistence}$(katm) | $P_{solid\ single\ phase}$(katm) | $P_{liquid\ single\ phase}$(katm) |
|---|---|---|---|
| 250 | 0.1258 | 0.070598 | 0.14087 |

As mentioned above differences in pressure are expected . The conditions shown in Table 16 ensure maintenance of the solid-liquid coexistence and therefore this ($T$,$P$) point and the corresponding single phase configurations

are used as the initial point for the Gibbs-Duhem integration. Solid and liquid configurations for the independent single phase simulations are extracted from the coexistence box according to the methods previously described. Independent single phase $N\sigma T$ and $NPT$ simulations for the solid and the liquid phase are carried out over $10^5$ time steps, 20% of which are used for equilibration, as described in the methodology section. For the integration a temperature increment of $\Delta T$=3K is applied. Results of the integration are shown in Table 17. Additionally, in Figures 27 and 28, the ($T$,$\varrho$) and ($P$-$T$) phase diagrams are shown compared to the available experimental data.
The ($T$,$\varrho$) phase diagram of the 5th structure and the OPLS force field appears to be shifted as observed in the other two systems previously investigated (the global minimum structure and CSP force field and the global minimum structure and OPLS force field). Once again, there is a clear underestimation of the triple point of at least 70 K. Difference between solid and liquid densities appears to be highly underestimated. Both solid and liquid regions appear to be overestimated although the trend appears to be reasonably correct.

Table 17: Single phase simulation results for the 5th structure and the OPLS force field

| $T$(K) | $\varrho_{solid}$(g/cm$^3$) | $\varrho_{liquid}$(g/cm$^3$) |
|---|---|---|
| 250 | 1.027783 | 0.957237 |
| 253 | 1.029451 | 0.961722 |
| 256 | 1.031027 | 0.965748 |
| 259 | 1.032942 | 0.967568 |
| 262 | 1.034217 | 0.968858 |
| 265 | 1.036087 | 0.970608 |
| 268 | 1.037757 | 0.972364 |
| 271 | 1.039477 | 0.9694548 |
| 274 | 1.041297 | 0.9709948 |
| 277 | 1.042967 | 0.9726148 |
| 280 | 1.044837 | 0.9742848 |

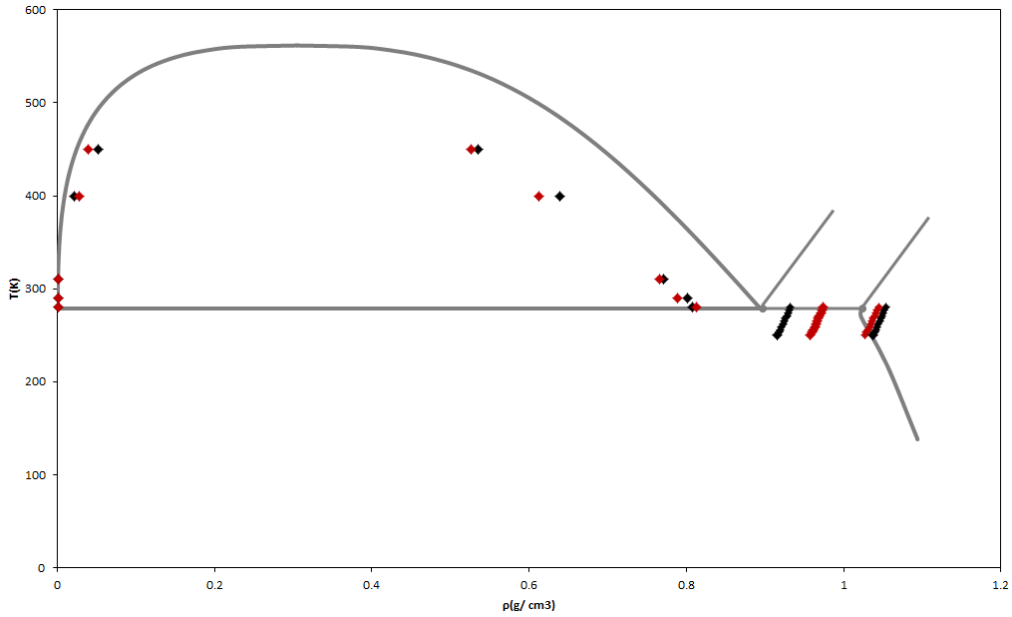


Figure 27: Phase diagram showing the boundaries of phase transition for the 5th and global minimum structure and OPLS force field. Simulation data for the VLE and SLE are represented with black diamonds for the 5th struture and with red diamonds for the global minimum structure. Experimental data are represented with grey lines and points [74] [20] [13] .

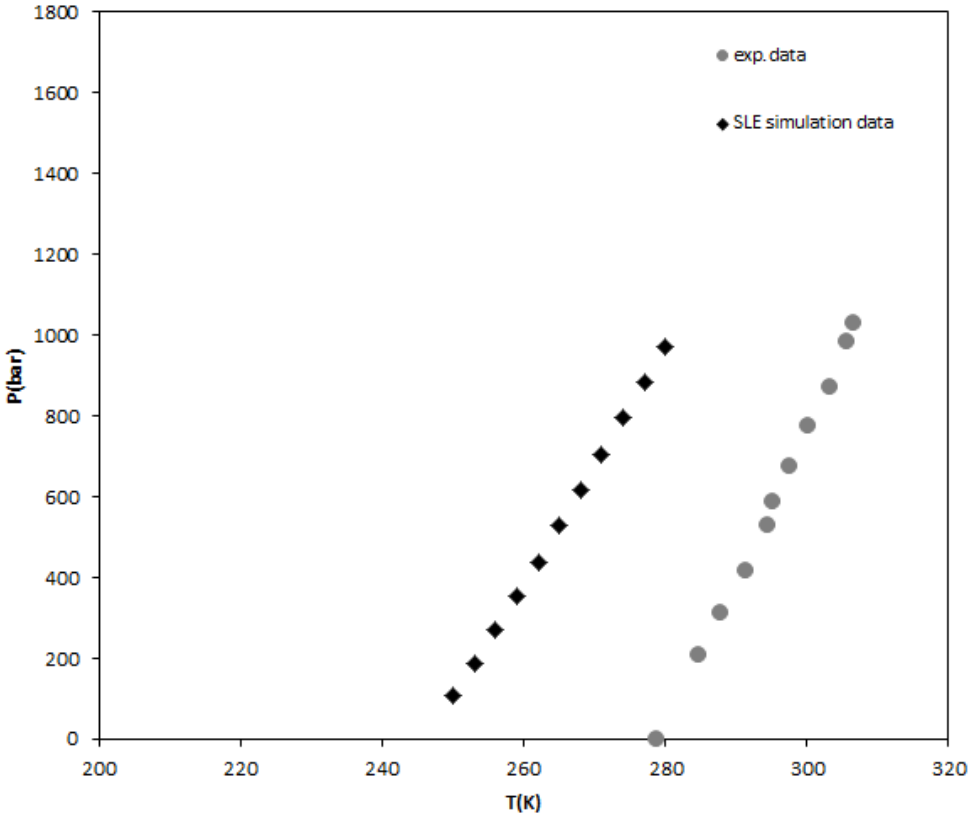


Figure 28: $P$-$T$ melting line for the 5th structure and the OPLS force field. Simulation data for the SLE region are represented with black diamonds and for the VLE region with red diamonds. Experimental data are represented with grey points [74][3].

The ($P$-$T$) diagram shows an underestimation of temperature for given pressure compared to the available data. This system results a melting point of approximately $T_m$=247 K at $P$=1 bar when the experimental value is $T_{m\,\mathrm{exp}}$=278.68 K at $P$=1 bar [6] [62].

# 5 Conclusions and future work

The melting point of organic compounds is a property of great significance and interest for many fields of research. Accurate prediction of it before a compound has been synthesized could accelerate the design of new materials. For this reason, development of a fully computational method is necessary and in this work the concept of such a method was investigated where prediction of the melting point of an organic compound was attempted by using as an input a computationally obtained crystal structure. Generally, as addressed in the literature review, the molecular simulation methods developed so far for the computation of melting points are not fully predictive, since they require an experimental crystal structure as input. This highlights the interest, originality and challenge of predicting of the melting point of a compound from first principles- given just the molecular diagram.

The application of our methodology to the Lennard-Jones potential can be characterized as successful, since the agreement between the available computational data and our predictions is very good. In the investigation conducted for benzene so far, two structurse computed by *CrystalPredictor* are used and two different force fields are employed. The resulting phase diagrams show an underestimation compared to the available data. The melting points computed at $P$=1 bar for the three systems (global minimum structure and CSP force field, global minimum structure and OPLS force field and 5th structure and OPLS force field) have difference of 30-80 K compared to the experimental values. Specifically, the application for the global minimum structure and CSP force field resulted a melting point of $T_m$=210 K at $P$=1 bar, which is 70K bellow the experimental value. In the case of the global minimum structure and OPLS force field the predicted $T_m$=200 K is underestimated by 80 K compared to the experimental value. The system of the 5th structure and OPLS force field has a better performance since the predicted melting point is $T_m$=247 K which is closer to the experimental value of $T_{m\,\mathrm{exp}}$=278.68 K. Accurate prediction depends on both the force field applied as well as the melting point computation method. Proper choice of force field affects the agreement between experimental and computational data.

To further examine the effect of the potential used in our methodology, we suggest to use a different electrostatic potential. The new potential de-

scribes the electrostatic interactions by using a distributed multipole expansion which offers a more accurate intermolecular potential [17] [14]. The distributed multipole model is used in both *CrystalOptimizer* and DMACRYS and is known to yield more accurate results that the exp-6 potential and point charges force field used in *CrystalPredictor*. However, to our knowledge, it has not yet been adopted in this context. Therefore, it will be very interesting to see how our methodology performs under this different potential. For this reason the most promising minima (those of the lowest lattice energy) identified during the global search for structures with *CrystalPredictor*, are further minimized using the DMACRYS software. This provides with a much more accurate [17], model which is based on a distributed multipole expansion of the electrostatic potential [64]. Although this is not reported here, the accuracy of this model has already verified in this work since in the resulting energy landscape the most stable polymorph for benzene (Form I) successfully corresponds to the most stable structure generated (i.e the global minimum). The force field necessary for the MD simulations is generated and some initial test runs under the $NVT$ ensemble have been carried out using DL_MULTI [59]. This software models rigid molecules whose intermolecular interactions include distributed multipoles. These multipoles can be extracted using the distributed multipole analysis (DMA) method) [17][65].

The distributed multipole expansion of the electrostatic potential provides higher accuracy but at the same time is computationally-demanding. Thus, it makes sense to explore other models as well, which can be less demanding. Our method is applied for the CSP exp-6 potential force field using the charge density obtained from more accurate level of theory of QM. There are levels of theory of QM for extracting charge density that have been proven to be more accurate compared to the HF we used for our simulations so far (such as MP2, PBE0, M06 and B3LYP). Thus, it is interesting to explore whether this increased accuracy would be translated into more accurate melting point predictions.

Finally, it is important to highlight the significance of an accurate computation of the melting point of an organic compound from first principles. In terms of CSP, the optimization algorithms applyed result in structures with a high level of accuracy. In terms of melting point prediction method, there is a lot of room for improvement. The method applied in this work

is a direct method, computationally demanding and the application of it for more realistic models- other than the LJ sphere- is very time consuming and has resulted a limited performance. To ensure reliable predictions, a more rigorous, free-energy based method, such as the PSCP method [19][21] presented in detail in the literature review, could be applied after a structure from CSP is obtained. This will can contribute to the development of a methodology for *ab initio* melting point prediction with high accuracy and general application.

# References


[1] P. M. Agrawal, B. M. Rice, and D. L. Thompson. Molecular dynamics study of the effects of voids and pressure in defect-nucleated melting simulations. *The Journal of Chemical Physics*, 118(21):9680–9688, 2003.

[2] P. M. Agrawal, B. M. Rice, and D. L. Thompson. Molecular dynamics study of the melting of nitromethane. *The Journal of Chemical Physics*, 119(18):9617–9627, 2003.

[3] J. Akella. Phase Diagram of Benzene to 35 kbar. *The Journal of Chemical Physics*, 55(2):793, 1971.

[4] S. Alavi and D. L. Thompson. Molecular dynamics studies of melting and some liquid-state properties of 1-ethyl-3-methylimidazolium hexafluorophosphate [emim][PF6]. *The Journal of chemical physics*, 122(15):154704, 2005.

[5] S. Alavi and D. L. Thompson. Simulations of melting of polyatomic solids and nanoparticles. *Molecular Simulation*, 32(12-13):999–1015, 2006.

[6] M. Azreg-Aïnou, A. Hüseynov, and B. Ibrahimolu. Phase equilibrium and metastability of liquid benzene at high pressures. *The Journal of Chemical Physics*, 124(20):204505, 2006.

[7] D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. a. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti, and I. K. Zhitkov. Towards crystal structure prediction of complex organic compounds–a report on the fifth blind test. *Acta Crystallographica. Section B, Structural science*, 67(Pt 6):535–51, 2011.

[8] A. Belonoshko. Molecular dynamics of MgSiO 3 perovskite at high pressures: Equation of state, structure, and melting transition. *Geochimica et Cosmochimica Acta*, 58(19):4039–4047, 1994.

[9] A. Belonoshko, R. Ahuja, and B. Johansson. Quasiab initio molecular dynamic study of Fe melting. *Physical review letters*, 1:17–20, 2000.

[10] A. Belonoshko and L. Dubrovinsky. Equations of state of MgSiO 3-perovskite and MgO (periclase) from computer simulations. *Physics of the earth and planetary ...*, 9201(96), 1996.

[11] I. Cacelli, G. Cinacchi, G. Prampolini, and A. Tani. Computer simulation of solid and liquid benzene with an atomistic interaction potential derived from ab initio calculations. *Journal of the American Chemical Society*, 126(43):14278–86, 2004.

[12] I. Cacelli, G. Cinacchi, G. Prampolini, and A. Tani. Modeling benzene with single-site potentials from ab initio calculations: a step toward hybrid models of complex molecules. *The Journal of Chemical Physics*, 120(8):3648–56, 2004.

[13] F. Cansell, D. Fabre, and J.-P. Petitet. Phase transitions and chemical transformations of benzene up to 550C and 30 GPa. *The Journal of Chemical Physics*, 99(10):7300, 1993.

[14] S. Cardamone, T. J. Hughes, and P. L. A. Popelier. Multipolar electrostatics. *Physical chemistry chemical physics : PCCP*, 16(22):10367–87, 2014.

[15] G. M. Day. Current approaches to predicting molecular organic crystal structures. *Crystallography Reviews*, 17(1):3–52, 2011.

[16] G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. a. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. a. Welch, H. a. Scheraga, Y. a. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf, and B. Schweizer. Significant

progress in predicting the crystal structures of small organic molecules–a report on the fourth blind test. *Acta Crystallographica. Section B, Structural science*, 65(Pt 2):107–25, 2009.

[17] G. M. Day, W. D. S. Motherwell, and W. Jones. Beyond the Isotropic Atom Model in Crystal Structure Prediction of Rigid Molecules: Atomic Multipoles versus Point Charges. *Crystal Growth & Design*, 5(3):1023–1033, 2005.

[18] D. M. Eike, J. F. Brennecke, and E. J. Maginn. Toward a robust and general molecular simulation method for computing solid-liquid coexistence. *The Journal of Chemical Physics*, 122(1):14115, 2005.

[19] D. M. Eike, J. F. Brennecke, and E. J. Maginn. Toward a robust and general molecular simulation method for computing solid-liquid coexistence. *The Journal of Chemical Physics*, 122(1):14115, 2005.

[20] D. M. Eike and E. J. Maginn. Atomistic simulation of solid-liquid coexistence for molecular systems: application to triazole and benzene. *The Journal of chemical physics*, 124(16):164503, 2006.

[21] D. M. Eike and E. J. Maginn. Atomistic simulation of solid-liquid coexistence for molecular systems: Application to triazole and benzene. *The Journal of Chemical Physics*, 124(16):164503, 2006.

[22] D. M. Eike and E. J. Maginn. Atomistic simulation of solid-liquid coexistence for molecular systems: application to triazole and benzene. *The Journal of chemical physics*, 124(16):164503, 2006.

[23] F. A. Escobedo. A unified methodological framework for the simulation of nonisothermal ensembles. *The Journal of chemical physics*, 123(4):044110, 2005.

[24] H. Feng, J. Zhou, and Y. Qian. Atomistic simulations of the solid-liquid transition of 1-ethyl-3-methyl imidazolium bromide ionic liquid. *The Journal of Chemical Physics*, 135(14):144501, 2011.

[25] D. Frenkel and S. Berend. *Understanding Molecular Simulation: from Algorithms to Applications*. San Diego, 2nd edition, 2002.

[26] D. Frenkel and A. J. C. Ladd. New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres. *The Journal of Chemical Physics*, 81(7):3188, 1984.

[27] D. Frenkel and A. J. C. Ladd. New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres. *The Journal of Chemical Physics*, 81(7):3188, 1984.

[28] G. Grochola. Constrained fluid lambda-integration: constructing a reversible thermodynamic path between the solid and liquid state. *The Journal of Chemical Physics*, 120(5):2122–6, 2004.

[29] G. Grochola. Further application of the constrained fluid lambda-integration method. *The Journal of Chemical Physics*, 122(4):46101, 2005.

[30] Y. Hiwatari, E. Stoll, and T. Schneider. Molecular-dynamics investigation of solid-liquid coexistence. *The Journal of Chemical Physics*, 68(8):3401, 1978.

[31] W. G. Hoover. Use of Computer Experiments to Locate the Melting Transition and Calculate the Entropy in the Solid Phase. *The Journal of Chemical Physics*, 47(12):4873, 1967.

[32] W. G. Hoover. Melting Transition and Communal Entropy for Hard Spheres. *The Journal of Chemical Physics*, 49(8):3609, 1968.

[33] S. Jayaraman and E. J. Maginn. Computing the melting point and thermodynamic stability of the orthorhombic and monoclinic crystalline polymorphs of the ionic liquid 1-n-butyl-3-methylimidazolium chloride. *The Journal of Chemical Physics*, 127(21):214504, 2007.

[34] J. K. Johnson, J. A. Zollweg, and K. E. Gubbins. The Lennard-Jones equation of state revisited. *Molecular Physics*, 78(3):591–618, 1993.

[35] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.*, 118(45):11225–11236, 1996.

[36] P. G. Karamertzanis and C. C. Pantelides. Ab initio crystal structure prediction-I. Rigid molecules. *Journal of Computational Chemistry*, 26(3):304–24, 2005.

[37] A. R. Katritzky, A. Lomaka, R. Petrukhin, R. Jain, M. Karelson, A. E. Visser, and R. D. Rogers. QSPR correlation of the melting point for pyridinium bromides, potential ionic liquids. *Journal of chemical information and computer sciences*, 42(1):71–4, 2002.

[38] A. V. Kazantsev, P. G. Karamertzanis, C. S. Adjiman, C. C. Pantelides, S. L. Price, P. T. A. Galek, G. M. Day, and A. J. Cruz-Cabeza. Successful prediction of a model pharmaceutical in the fifth blind test of crystal structure prediction. *International Journal of Pharmaceutics*, 418(2):168–78, 2011.

[39] J. G. Kirkwood. Statistical Mechanics of Fluid Mixtures. *The Journal of Chemical Physics*, 3(5):300, 1935.

[40] D. A. Kofke. Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line. *The Journal of Chemical Physics*, 98(5):4149, 1993.

[41] D. A. Kofke. Gibbs-Duhem integration: a new method for direct evaluation of phase coexistence by molecular simulation. *Molecular Physics*, 78(6):1331–1336, 1993.

[42] B. B. Laird. Direct Calculation of the Crystal-Melt Interfacial Free Energy via Molecular Dynamics. 35(2):17802–17812, 2005.

[43] J. E. Lennard-Jones and A. F. Devonshire. Critical and Co-Operative Phenomena. III. A Theory of Melting and the Structure of Liquids. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 169(938):317–338, 1939.

[44] Z.-L. Liu, L.-C. Cai, X.-R. Chen, and F.-Q. Jing. Molecular dynamics simulations of the melting curve of tantalum under pressure. *Physical Review B*, 77(2):024103, 2008.

[45] J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, W. T. M. Mooij, S. L. Price, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer,

and D. E. Williams. A test of crystal structure prediction of small organic molecules. *Acta Crystallographica Section B Structural Science*, 56(4):697–714, 2000.

[46] S.-N. Luo, T. Ahrens, T. Çan, A. Strachan, W. Goddard, and D. Swift. Maximum superheating and undercooling: Systematics, molecular dynamics simulations, and dynamic experiments. *Physical Review B*, 68(13):134206, 2003.

[47] S.-N. Luo and T. J. Ahrens. Superheating systematics of crystalline solids. *Applied Physics Letters*, 82(12):1836–1838, 2003.

[48] S.-N. Luo, A. Strachan, and D. C. Swift. Nonequilibrium melting and crystallization of a model Lennard-Jones system. *The Journal of chemical physics*, 120(24):11640–9, 2004.

[49] S.-N. Luo, A. Strachan, and D. C. Swift. Nonequilibrium melting and crystallization of a model Lennard-Jones system. *The Journal of chemical physics*, 120(24):11640–9, 2004.

[50] J. F. Lutsko, D. Wolf, S. R. Phillpot, and S. Yip. Molecular-dynamics study of lattice-defect-nucleated melting in metals using an embedded-atom-method potential. *Physical Review B*, 40(5):2841–2855, 1989.

[51] C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek, and P. A. Wood. Mercury CSD 2.0 new features for the visualization and investigation of crystal structures. *Journal of Applied Crystallography*, 41(2):466–470, 2008.

[52] E. A. Mastny and J. J. de Pablo. Melting line of the Lennard-Jones system, infinite size, and full potential. *The Journal of chemical physics*, 127(10):104504, 2007.

[53] J. R. Morris, C. Z. Wang, K. M. Ho, and C. T. Chan. Melting line of aluminum from simulations of coexisting phases. *Phys. Rev. B*, 49(5):3109–3115, 1994.

[54] W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. Scheraga, B. Schweizer,

M. U. Schmidt, B. P. van Eijck, P. Verwer, and D. E. Williams. Crystal structure prediction of small organic molecules: a second blind test. *Acta Crystallographica Section B Structural Science*, 58(4):647–661, 2002.

[55] M. Nayhouse, A. M. Amlani, and G. Orkoulas. A Monte Carlo study of the freezing transition of hard spheres. *Journal of physics. Condensed matter : an Institute of Physics journal*, 23(32):325106, 2011.

[56] G. Orkoulas and M. Nayhouse. Communication: A simple method for simulation of freezing transitions. *The Journal of chemical physics*, 134(17):171104, 2011.

[57] G. Orkoulas and M. Nayhouse. Communication: A simple method for simulation of freezing transitions. *The Journal of chemical physics*, 134(17):171104, 2011.

[58] S. Phillpot, J. Lutsko, D. Wolf, and S. Yip. Molecular-dynamics study of lattice-defect-nucleated melting in silicon. *Physical Review B*, 40(5):2831–2840, 1989.

[59] S. L. Price, S. Hamad, a. Torrisi, P. G. Karamertzanis, M. Leslie, and C. R. a. Catlow. Applications Of Dl_poly And Dl_multi To Organic Molecular Crystals. *Molecular Simulation*, 32(12-13):985–997, 2006.

[60] P. Raiteri, R. Martonák, and M. Parrinello. Exploring polymorphism: the case of benzene. *Angewandte Chemie (International ed. in English)*, 44(24):3769–73, 2005.

[61] N. S. Ramrattan. Simulation and Theoretical Perspectives of the Phase Behaviour of Solids , Liquids and Gases using the Mie Family of Intermolecular Potentials by. 2014.

[62] Robert D. Goodwin. Goodwin.pdf. *The Journal of Physical Chemistry C*, 17(4):1542–1636, 1988.

[63] J. Solca, A. J. Dyson, G. Steinebrunner, B. Kirchner, and H. Huber. Melting curves for neon calculated from pure theory. *The Journal of Chemical Physics*, 108(10):4107, 1998.

[64] A. J. Stone. Distributed Multipole Analysis: Stability for Large Basis Sets. *Journal of Chemical Theory and Computation*, 1(6):1128–1132, 2005.

[65] a. J. Stone and M. Alderton. Distributed multipole analysis Methods and applications. *Molecular Physics*, 100(1):221–233, 2002.

[66] N. Sun, X. He, K. Dong, X. Zhang, X. Lu, H. He, and S. Zhang. Prediction of the melting points for two kinds of room temperature ionic liquids. *Fluid Phase Equilibria*, 246(1-2):137–142, 2006.

[67] C. Vega, E. Sanz, J. L. F. Abascal, and E. G. Noya. Determination of phase diagrams via computer simulation: methodology and applications to water, electrolytes and proteins. *Journal of Physics: Condensed Matter*, 20(15):153101, 2008.

[68] G. F. Velardez, S. Alavi, and D. L. Thompson. Molecular dynamics studies of melting and solid-state transitions of ammonium nitrate. *The Journal of Chemical Physics*, 120(19):9151–9, 2004.

[69] S. W. Watt, J. A. Chisholm, W. Jones, and S. Motherwell. A molecular dynamics simulation of the melting points and glass transition temperatures of myo- and neo-inositol. *The Journal of Chemical Physics*, 121(19):9565–73, 2004.

[70] N. B. Wilding and A. D. Bruce. Freezing by Monte Carlo Phase Switch. *Phys. Rev. Lett.*, 85(24):5138–5141, 2000.

[71] S. Yoo, X. C. Zeng, and S. S. Xantheas. On the phase diagram of water with density functional theory potentials: The melting temperature of ice I[sub h] with the Perdew–Burke–Ernzerhof and Becke–Lee–Yang–Parr functionals. *The Journal of Chemical Physics*, 130(22):221102, 2009.

[72] Y. Zhang and E. J. Maginn. A comparison of methods for melting point calculation using molecular dynamics simulations. *The Journal of Chemical Physics*, 136(14):144116, 2012.

[73] Y. Zhang and E. J. Maginn. Toward Fully in Silico Melting Point Prediction Using Molecular Simulations. *Journal of Chemical Theory and Computation*, 9(3):1592–1599, 2013.

[74] X. S. Zhao, B. Chen, S. Karaborni, and J. I. Siepmann. Vapor-liquid and vapor-solid phase equilibria for united-atom benzene models near their

triple points: the importance of quadrupolar interactions. *The journal of physical chemistry. B*, 109(11):5368–74, 2005.

[75] L. Zheng, S.-N. Luo, and D. L. Thompson. Molecular dynamics simulations of melting and the glass transition of nitromethane. *The Journal of Chemical Physics*, 124(15):154504, 2006.

# A Appendix

Table 18: Simulation results for the Lennard-Jones system for $T^*$=0.83-1.38.

| $T^*$ | $\varrho^*_{solid}$ | $\varrho^*_{liquid}$ |
|---|---|---|
| 0.83142 | 0.980680453 | 0.882804234 |
| 0.85449 | 0.984123158 | 0.888056736 |
| 0.87756 | 0.987599378 | 0.893171801 |
| 0.90063 | 0.991101469 | 0.898168708 |
| 0.9237 | 0.994620957 | 0.903059996 |
| 0.94677 | 0.998149407 | 0.907853855 |
| 0.96984 | 1.001678913 | 0.91255573 |
| 0.99291 | 1.005202358 | 0.917169384 |
| 1.01598 | 1.008713525 | 0.921697586 |
| 1.03905 | 1.012207109 | 0.926142566 |
| 1.06212 | 1.015678684 | 0.930506294 |
| 1.08101 | 1.017586633 | 0.939669354 |
| 1.10826 | 1.022542069 | 0.938997515 |
| 1.13133 | 1.025928745 | 0.943128836 |
| 1.1544 | 1.029282974 | 0.947186633 |
| 1.17747 | 1.032603548 | 0.951173 |
| 1.20054 | 1.035889664 | 0.955090104 |
| 1.22361 | 1.03914086 | 0.958940158 |
| 1.24668 | 1.042356958 | 0.962725412 |
| 1.26975 | 1.045538014 | 0.966448131 |
| 1.29282 | 1.048684272 | 0.970110577 |
| 1.31589 | 1.051796131 | 0.973714996 |
| 1.33896 | 1.054874108 | 0.977263602 |
| 1.36203 | 1.057918817 | 0.980758566 |
| 1.3851 | 1.060930942 | 0.984202008 |

Table 19: Simulation results for the Lennard-Jones system for $T^*$=1.4-2.16.

| $T^*$ | $\varrho^*_{solid}$ | $\varrho^*_{liquid}$ |
|---|---|---|
| 1.40817 | 1.063911218 | 0.987595989 |
| 1.43124 | 1.066860418 | 0.990942504 |
| 1.45431 | 1.069779339 | 0.994243479 |
| 1.47738 | 1.07266879 | 0.997500768 |
| 1.50045 | 1.075529583 | 1.000716151 |
| 1.52352 | 1.078362529 | 1.003891332 |
| 1.54659 | 1.081168429 | 1.007027941 |
| 1.56966 | 1.08394807 | 1.010127536 |
| 1.59273 | 1.086702224 | 1.013191599 |
| 1.6158 | 1.089431642 | 1.016221542 |
| 1.63887 | 1.092137054 | 1.01921871 |
| 1.66194 | 1.094819167 | 1.022184377 |
| 1.68501 | 1.097478665 | 1.025119755 |
| 1.70808 | 1.100116208 | 1.028025994 |
| 1.73115 | 1.102732429 | 1.030904183 |
| 1.75422 | 1.10532794 | 1.033755356 |
| 1.77729 | 1.107903326 | 1.036580491 |
| 1.80036 | 1.110459148 | 1.039380514 |
| 1.82343 | 1.112995947 | 1.042156304 |
| 1.8465 | 1.115514237 | 1.044908692 |
| 1.86957 | 1.118014513 | 1.047638464 |
| 1.89264 | 1.120497247 | 1.050346367 |
| 1.91571 | 1.122962891 | 1.053033104 |
| 1.93878 | 1.125411878 | 1.055699345 |
| 1.96185 | 1.127844621 | 1.058345721 |
| 1.98492 | 1.130261514 | 1.060972832 |
| 2.00799 | 1.132662937 | 1.063581246 |
| 2.03106 | 1.135049249 | 1.066171501 |
| 2.05413 | 1.137420797 | 1.068744106 |
| 2.0772 | 1.139777911 | 1.071299546 |
| 2.10027 | 1.142120907 | 1.073838278 |
| 2.12334 | 1.144450087 | 1.076360739 |
| 2.14641 | 1.146765741 | 1.078867342 |
| 2.16948 | 1.149068147 | 1.08135848 |

Table 20: Simulation results for the Lennard-Jones system for $T^*$=2.19-2.97.

| $T^*$ | $\varrho^*_{solid}$ | $\varrho^*_{liquid}$ |
|---|---|---|
| 2.19255 | 1.151357569 | 1.083834524 |
| 2.21562 | 1.153634262 | 1.086295831 |
| 2.23869 | 1.15589847 | 1.088742737 |
| 2.26176 | 1.158150426 | 1.091175564 |
| 2.28483 | 1.160390356 | 1.093594616 |
| 2.3079 | 1.162618475 | 1.096000186 |
| 2.33097 | 1.164834989 | 1.098392551 |
| 2.35404 | 1.167040098 | 1.100771976 |
| 2.37711 | 1.169233992 | 1.103138714 |
| 2.40018 | 1.171416855 | 1.105493006 |
| 2.42325 | 1.173588865 | 1.107835083 |
| 2.44632 | 1.175750192 | 1.110165165 |
| 2.46939 | 1.177901001 | 1.112483464 |
| 2.49246 | 1.180041448 | 1.114790182 |
| 2.51553 | 1.182171689 | 1.117085513 |
| 2.5386 | 1.184291869 | 1.119369642 |
| 2.56167 | 1.186402131 | 1.121642748 |
| 2.58474 | 1.188502614 | 1.123905001 |
| 2.60781 | 1.190593449 | 1.126156568 |
| 2.63088 | 1.192674767 | 1.128397604 |
| 2.65395 | 1.194746692 | 1.130628264 |
| 2.67702 | 1.196809343 | 1.132848693 |
| 2.70009 | 1.19886284 | 1.135059033 |
| 2.72316 | 1.200907295 | 1.137259419 |
| 2.74623 | 1.202942818 | 1.139449984 |
| 2.7693 | 1.204969516 | 1.141630854 |
| 2.79237 | 1.206987494 | 1.143802152 |
| 2.81544 | 1.208996853 | 1.145963996 |
| 2.83851 | 1.21099769 | 1.148116501 |
| 2.86158 | 1.212990101 | 1.150259778 |
| 2.88465 | 1.214974179 | 1.152393936 |
| 2.90772 | 1.216950016 | 1.154519077 |
| 2.93079 | 1.218917699 | 1.156635305 |
| 2.95386 | 1.220877315 | 1.158742716 |
| 2.97693 | 1.222828947 | 1.160841407 |

Table 21: Single phase simulation details for the Lennard-Jones system.

| | |
|---|---|
| Ensemble for solid phase | $N\sigma T$ |
| Ensemble for liquid phase | $NPT$ |
| Number of molecules | 2916 |
| cutoff | $4\sigma$ |
| Simulation time | $10^5$ time steps |
| Equilibration time | $10^4$ time steps |

Table 22: Solid-liquid coexistence simulation details for the Lennard-Jones system.

| | |
|---|---|
| Ensemble for solid phase | $NVT$ |
| Number of molecules | 11594 |
| cutoff | $4\sigma$ |
| Simulation time | $10^6$ time steps |
| Equilibration time | $10^5$ time steps |

Table 23: Simulation results of the heating isobar at $P$=1 bar for the global minimum structure and CSP force field.

| $T$(K) | $\varrho_{solid}$(g/cm$^3$) |
|---|---|
| 2 | 1.079495 |
| 10 | 1.078195 |
| 30 | 1.071705 |
| 50 | 1.064707 |
| 70 | 1.058062 |
| 90 | 1.050889 |
| 110 | 1.042402 |
| 130 | 1.034401 |
| 150 | 1.025555 |
| 170 | 1.016015 |
| 190 | 1.005677 |
| 210 | 0.994268 |
| 230 | 0.981694 |
| 250 | 0.96731 |
| 270 | 0.950406 |
| 290 | 0.929691 |
| 310 | 0.920969 |

Table 24: Simulation results of the cooling isobar at $P$=1 bar for the global minimum structure and CSP force field.

| $T$(K) | $\varrho_{liquid}$(g/cm$^3$) |
|---|---|
| 310 | 0.794063 |
| 260 | 0.848902 |
| 240 | 0.870441 |
| 220 | 0.892171 |
| 200 | 0.914716 |
| 180 | 0.935875 |
| 160 | 0.954498 |
| 140 | 0.971947 |
| 120 | 0.984481 |
| 100 | 0.996798 |

Table 25: Simulation $T$-$P$ results for Gibbs-Duhem integration for the global minimum structure and CSP force field.

| $T$(K) | $P$(katm) |
|---|---|
| 240 | 0.955375 |
| 243 | 1.034462975 |
| 246 | 1.118040629 |
| 249 | 1.207578166 |
| 252 | 1.29523966 |
| 255 | 1.374327634 |
| 258 | 1.463865171 |
| 261 | 1.552979237 |
| 264 | 1.642328461 |
| 267 | 1.732314932 |
| 270 | 1.821358811 |

Table 26: Simulation results of the heating isobar at $P$=1 bar for the global minimum structure and OPLS force field.

| $T$(K) | $\varrho_{solid}$(g/cm$^3$) |
|---|---|
| 2 | 1.179259 |
| 10 | 1.176186 |
| 138 | 1.124342 |
| 218 | 1.083002 |
| 270 | 1.048089 |
| 290 | 1.031799 |
| 320 | 1.001416 |
| 340 | 0.971318 |
| 360 | 0.814029 |

Table 27: Simulation results of the cooling isobar at $P$=1 bar for the global minimum structure and OPLS force field.

| $T$(K) | $\varrho_{liquid}$(g/cm$^3$) |
|---|---|
| 312 | 0.872911 |
| 300 | 0.888875 |
| 281 | 0.911944 |
| 270 | 0.925306 |
| 240 | 0.963207 |
| 210 | 0.998351 |
| 180 | 1.025883 |

Table 28: Simulation $T$-$P$ results for Gibbs-Duhem integration for the global minimum structure and OPLS force field.

| $T$(K) | $P$(katm) |
|---|---|
| 250 | 1.525237975 |
| 253 | 1.606376167 |
| 256 | 1.691904294 |
| 259 | 1.779565787 |
| 262 | 1.868379852 |
| 265 | 1.957629077 |
| 268 | 2.047415548 |
| 271 | 2.136010743 |
| 274 | 2.22349848 |
| 277 | 2.309026607 |
| 280 | 2.393188462 |

Table 29: Simulation results of the heating isobar at $P$=1 bar for the 5th structure and OPLS force field.

| $T$(K) | $\varrho_{solid}$(g/cm$^3$) |
|---|---|
| 2 | 1.047652 |
| 10 | 1.044985 |
| 110 | 1.007625 |
| 130 | 0.99899 |
| 138 | 0.995422 |
| 150 | 0.989945 |
| 170 | 1.026411 |
| 190 | 1.032416 |
| 218 | 1.038481 |
| 270 | 1.007095 |
| 300 | 0.894866 |

Table 30: Simulation results of the cooling isobar at $P$=1 bar for the 5th structure and OPLS force field.

| $T$(K) | $\varrho_{liquid}$(g/cm$^3$) |
|---|---|
| 312 | 0.872947 |
| 281 | 0.911937 |
| 280 | 0.913056 |
| 250 | 0.950686 |
| 220 | 0.986189 |

Table 31: Simulation $T$-$P$ results for Gibbs-Duhem integration for the 5th structure and OPLS force field.

| $T$(K) | $P$(katm) |
|---|---|
| 250 | 0.105734 |
| 253 | 0.184521975 |
| 256 | 0.265160167 |
| 259 | 0.348188294 |
| 262 | 0.433125831 |
| 265 | 0.520213568 |
| 268 | 0.607875061 |
| 271 | 0.696389126 |
| 274 | 0.785575598 |
| 277 | 0.870513135 |
| 280 | 0.95860833 |