Supporting Information for:

A State Between Liquid and Crystal: Locally Crystalline but with the Structure Factor of a Liquid

James P. Mithen, and Richard P. Sear*

Department of Physics, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom

E-mail: r.sear@surrey.ac.uk

*To whom correspondence should be addressed
This supporting information has four sections: Computer simulation details, Lechner-Dellago order parameters for crystalline ordering, calculation of the structure factor $S(k)$, and crystallisation of the amorphous state.

**Computer Simulations Details**

**Monte Carlo Simulations**

The time evolution in our simulations is given by $NPT$ Monte-Carlo (MC). The simulations are at constant number of particles, $N$, pressure, $P$, and temperature, $T$. We use the normal Metropolis scheme with a maximum trial displacement, and a maximum volume change parameter, that ensure approximately 40% move acceptance; the exact values used depend on $P$ and $T$. All simulations are carried out in a cubic box. Our use of $NPT$ MC is similar to that in earlier nucleation studies.\textsuperscript{1,2} Time in our simulations is measured in units of ‘cycles’, with a single MC cycle, denoted by $\tau$, corresponding to (on average) a single attempted displacement move per molecule and a single attempted volume change, accepted or rejected in the normal way.\textsuperscript{3}

All results in the main text are for systems of 5000 molecules. However, we studied the formation of the amorphous state in system sizes up to 100,000 molecules, and found that this did not change the amorphous states we found.

The amorphous configurations in the main text formed via nucleation in the liquid phase, followed by growth, which we studied using the Forward Flux Sampling algorithm.\textsuperscript{4,5} We studied $\approx 150$ independent systems. Results for this step were presented in earlier work.\textsuperscript{6} In that work we studied the critical nuclei at the top of the barrier, and found that they were mixtures of all three polymorphs. Combining those results with results here, we see that our amorphous state is a mixture of polymorphs from nucleation onwards. It is worth noting that these mixed nuclei formed at one set of conditions. Under other conditions, we did observe nucleation and growth of more conventional nuclei, in which one polymorph dominated.\textsuperscript{6}

**Molecular Dynamics Simulations to calculate $D$**

Using NVE molecular dynamics, we computed diffusion coefficients. The diffusion coefficient $D \approx 4.6 \times 10^{-5} \sigma^2 / \tau_{MD}$ in the amorphous state, compared to $\approx 3.6 \times 10^{-5} \sigma^2 / \tau_{MD}$ for the pure fcc crystal and $\approx 1.3 \times 10^{-3} \sigma^2 / \tau_{MD}$ for the metastable liquid. Here the unit of time in our molecular dynamics simulations, is $\tau_{MD} = 0.1 \sqrt{m \sigma^2 / \varepsilon}$, where $m$ is the mass of the particles.

**Lechner-Dellago Order Parameters for Crystalline Ordering**

To identify crystalline molecules in our simulations we use the now standard method of Lechner and Dellago (henceforth LD).\textsuperscript{7} In contrast to other approaches, this method is designed specifically to discriminate between hcp, bcc and fcc crystal structures, and is therefore ideal for studying systems that exhibit multiple crystal polymorphs. By assessing the local environment of each molecule in the simulation box, the method allows each molecule to be identified as fcc-like, bcc-like, hcp-like, or liquid-like. This is done as follows.
First, for every molecule \(i\) in the simulation, we compute the complex vector

\[
q_{lm}(i) = \frac{1}{N_b(i)} \sum_{j=1}^{N_b(i)} Y_{lm}(\hat{r}_{ij}) .
\]

(1)

Here the functions \(Y_{lm}\) are spherical harmonics, \(-l \leq m \leq l\), \(N_b(i)\) is the number of neighbours of molecule \(i\), and \(\hat{r}_{ij}\) is a unit vector connecting molecules \(i\) and \(j\). The summation is over all molecules that are neighbours to particle \(i\). The vector is computed for both \(l = 4\) and \(l = 6\).

In our implementation, rather than using a distance cutoff, we define the neighbours of molecule \(i\) to be the 12 nearest molecules. Thus \(N_b(i) = 12\) for every molecule. The advantages of this neighbourhood definition are discussed in detail by Mickel et al.\(^8\)

Next, for each molecule \(i\) we take an average of \(q_{lm}(i)\) over its local neighbourhood by computing the complex vector

\[
\bar{q}_{lm}(i) = \frac{1}{N_b(i) + 1} \sum_{j=0}^{N_b(i)} q_{lm}(j) .
\]

(2)

Here the sum starting at \(j = 0\) runs over all neighbours of molecule \(i\) plus the molecule itself. This averaging procedure — which effectively means that the second shell of a molecules’ neighbours is taken into account — is the key step that allows the LD method to distinguish effectively between crystalline polymorphs. See Lechner and Dellago\(^7\) for further discussion. Using the \(\bar{q}_{lm}(i)\) vectors, we can compute for every molecule \(i\) the local bond order parameters

\[
\bar{q}_l(i) = \sqrt{\frac{4\pi}{2l + 1} \sum_{m=-l}^{l} |\bar{q}_{lm}(i)|^2} ,
\]

(3)

and

\[
\bar{w}_l(i) = \frac{\sum_{m_1 + m_2 + m_3 = 0} (l \quad l \quad l) \quad (m_1 \quad m_2 \quad m_3) \quad \bar{q}_{l m_1}(i) \quad \bar{q}_{l m_2}(i) \quad \bar{q}_{l m_3}(i)}{\left( \sum_{m=-l}^{l} |\bar{q}_{lm}(i)|^2 \right)^{3/2}} .
\]

(4)

The term in parentheses on the top line of Eq. (4) is the Wigner 3\( - j \) symbol.\(^9\) We use these expressions for \(l = 4\) and \(l = 6\), to compute three numbers for each molecule: \(\bar{q}_6\), \(\bar{w}_4\) and \(\bar{w}_6\) (\(\bar{q}_4\) is not needed in the classification procedure).

**Classification of molecules into liquid-like, fcc-like, hcp-like and bcc-like**

As shown in the left-hand column of Figure S1, the values of \(\bar{q}_6(i)\), \(\bar{w}_6(i)\) and \(\bar{w}_4(i)\) allow us to classify molecule \(i\) as one of liquid-like, fcc-like, hcp-like or bcc-like. With reference to the left-hand column of Figure S1, we achieve this as follows. Firstly, if \(\bar{q}_6(i) < 0.3\) (top left panel), molecule \(i\) is classified as liquid-like. If \(\bar{q}_6(i) > 0.3\) it is classified as crystalline. If it is crystalline and if \(\bar{w}_6(i) > 0\), it is classified as bcc-like (middle left panel). Finally, if the molecule is crystalline but not bcc-like, we look at \(\bar{w}_4\) (bottom left panel). If \(\bar{w}_4 > 0\), the molecule is hcp-like, otherwise
Figure S1: Probability density functions (pdfs) for our three local bond order parameters $\bar{q}_6$ (top row), $\bar{w}_6$ (middle row), and $\bar{w}_4$ (bottom row). The left-hand column shows the results for bulk liquids (blue), fcc crystals (yellow), hcp crystals (orange), and bcc crystals (green). These pdfs for bulk phases were computed from single configurations of around 5000 particles, at $T = 0.002, P = 0.055$. All phases were well equilibrated, and the crystals were all perfect defect-free lattices. A window filter was used to smooth the curves. The right-hand column shows the results for the configuration number 1 of our amorphous state. This is the one shown in Figure 3, which ultimately crystallises into the fcc polymorph. This configuration has 5000 molecules. We did not smooth these curves. The blue, yellow, orange and green curves are for molecules identified using the order parameters as liquid, fcc, hcp, and bcc, respectively. In all panels the vertical dashed lines denote the values we use for classification.

it is fcc-like.

Note that for example, if we look at the top-left panel, all but a tiny fraction of molecules in a liquid have $\bar{q}_6 < 0.3$, while all but a tiny fraction of the molecules in any of the three crystal polymorphs have $\bar{q}_6 > 0.3$. It is this lack of overlap between the liquid and crystal probability density functions for $\bar{q}_6$, that makes $\bar{q}_6$ a good order parameter. And the same applies for $\bar{w}_6$ and $\bar{w}_4$. 
Figure S2: Probability density functions (pdfs) for our three local bond order parameters $\bar{q}_6$ (top row), $\bar{w}_6$ (middle row), and $\bar{w}_4$ (bottom row). The left-hand column shows the same results for bulk liquids (blue), fcc crystals (yellow), hcp crystals (orange), and bcc crystals (green), as in Figure S1. See the caption to Figure S1 for details. The right-hand column shows the results for the configuration number 5 of our amorphous state. This is the one shown in Figure 4, which ultimately crystallises into the bcc polymorph. This configuration has 5000 molecules. We did not smooth these curves. The blue, yellow, orange and green curves are for molecules identified using the order parameters as liquid, fcc, hcp, and bcc, respectively. In all panels the vertical dashed lines denote the values we use for classification.

### Crystalline clusters

Once each molecule has been classified, we can determine how many molecules are in a given polymorph, how big are the clusters of crystalline molecules, and how big are the clusters of molecules that are all in a particular polymorph.

For crystalline clusters, we consider only those molecules identified as crystalline. A molecule is considered part of a crystalline cluster if it is within a cut-off distance $r_{cl}$ of any other molecule in the cluster. We set $r_{cl}$ to be the average of the first and second neighbour distance for a perfect fcc crystal, which has a known density as determined from NPT simulations.

We also do this for connected clusters of molecules all in the same polymorph. A cluster of polymorph fcc, for example, is the set of molecules identified as fcc, in which every one of which
is within $r_{cl}$ of at least one other molecule of the cluster.

**Order parameter probability density functions in the amorphous state**

In the right-hand column of Figure S1, we have plotted the amorphous-state probability density functions of our order parameters. These probability density functions can be compared to those in a bulk liquid, and in bulk defect-free crystal lattices, which are in the left-hand column. Note that in the left-hand column the blue, yellow, orange and green curves are for what we know are a bulk liquid, and fcc, hcp and bcc crystals, respectively, while in the right-hand column the curves are for molecules classified using these order parameters as liquid-like, fcc-like, hcp-like and bcc-like, respectively.

The first thing to note is that the few (150 out of 5000 in total) molecules identified as liquid-like are all in environments that are rare in the bulk liquid. Most molecules in the liquid have values of $\bar{q}_6$ in the range 0.1 to 0.2, whereas they are in the range 0.2 to 0.3 in the amorphous state. Thus after transformation to the amorphous state, there are very few molecules left in bulk-liquid-like environments.

It is also true that many molecules identified as fcc, hcp or bcc are in environments that are not typical of the defect-free bulk crystal lattices, in the sense of having values of $\bar{q}_6$, $\bar{w}_6$ and $\bar{w}_4$ that are not within a standard deviation of the mode of the probability density functions in the bulk crystals. If we look at the top row, we see that in the amorphous state the values of $\bar{q}_6$ are typically lower than in the bulk crystal, i.e., biased towards the values in the liquid state. In this sense the local crystalline ordering is weaker, on average, in the amorphous state than in defect-free bulk crystals. Also, if we look at the bottom two rows of Figure S1, we see that many molecules in the amorphous state have relatively poorly defined polymorphic state. In the middle row we can see that some molecules have ordering that is not unambiguously that of a close packed (fcc or hcp) lattice or that in the bcc polymorph, but is intermediate. In the bottom row, the picture is similar, there are molecules that are, according to the $\bar{w}_4$ order parameter, intermediate between the hcp and fcc polymorphs.

It is not surprising that many molecules have values of the LD order parameters that are intermediate between the values characteristic of the four bulk phases. The domains of the three polymorphs are such that many molecules are either at an interface with another polymorph or only one or two molecular diameters away from an interface. This is clear from the snapshots of the individual polymorphs in Figure S3. Recall that the LD order parameters for a molecule contains information on both the neighbours and the next neighbours.

The configuration whose LD order parameters are shown in the right-hand column of Figure S1 crystallises into the fcc polymorph. For comparison, in Figure S2, we have plotted the probability density functions of LD order parameters of a configuration that crystallises into the bcc polymorph. The two sets of probability density functions are similar.

**Calculation of the Structure Factor $S(k)$**

The static structure factor of a configuration of $N$ particles is defined by

$$S(k) = \frac{1}{N} \langle |n(k)|^2 \rangle_k,$$  \hspace{1cm} (5)
where $k$ is the wave-vector, and $\lvert k \rvert$ is its magnitude. Here, $\langle \cdots \rangle_k$ denotes an average over wave-vectors of equal magnitude. The particle density

$$n(k) = \sum_{j=1}^{N} \exp(-ik \cdot r_j),$$

(6)

with $r_j$ the position of particle $j$.

We compute $S(k)$ for a single configuration of particles directly from the definition in Eq. (5). To do this, we first compute the allowed wave-vectors in the cubic simulation box of length $L$. These are given by

$$k = \frac{2\pi}{L} (n_x, n_y, n_z),$$

(7)

for integers $n_x$, $n_y$ and $n_z$. For each allowed wave-vector we then compute $n(k)$ using Eq. (6). Finally, we average over wave-vectors of equal magnitude. This procedure gives a value for $S(k)$ for every value of $k$ compatible with the periodic boundary conditions.

Figure S3: The three $S(k)$’s calculated for each individual polymorph. For example, the fcc $S(k)$ is calculated using only molecules identified as fcc, excluding bcc-like, hcp-like and liquid-like molecules. This is for a typical amorphous configuration that is identified as 97% crystalline, with 61% fcc, 24% bcc and 12% hcp. To the right of each $S(k)$ plot we show the molecular positions it was computed from. These $S(k)$’s have not been smoothed.
**Smoothed $S(k)$**

Often the $S(k)$ resulting from this procedure is noisy. Therefore, where appropriate (typically when there are no Bragg peaks in the $S(k)$ profile), we smooth $S(k)$ using a simple window filter. This filter simply assigns to each allowed value of $k$, an $S(k)$ based on the average of nearby $S(k)$ values. We average over 15 values of $S(k)$, i.e., for each $k$ value we take the 7 nearest $S(k)$ values (with $k$ compatible with the box) on either side as well as the central value. All $S(k)$ plots appearing in the main paper with the exception of those that exhibit Bragg peaks are smoothed in this way.

**Partial $S(k)$**

In the main text we give $S(k)$’s calculated for all the molecules in an amorphous configuration. For comparison, in Figure S3 we have plotted the $S(k)$’s calculated using only the molecules identified as fcc, only those identified as bcc, and only those identified as hcp. The $S(k)$’s are a little noisier, as by separating out the polymorphs, each curve is calculated for fewer molecules. The maximum peak height for the fcc and bcc polymorphs, is higher than that of the $S(k)$ calculated for all molecules (maximum of $\approx 12$ as compared to $\approx 8$) but the $S(k)$’s are very different from those of bulk crystals (shown in the main text). Here a majority of the molecules are locally fcc and the fcc cluster percolates, see top snapshot of Figure S3, but the $S(k)$ of this cluster is still far from that of a bulk fcc crystal.

**Comparison with $S(k)$’s obtained from the computer simulation of the crystallisation of water**

The Gaussian Core Model is not the only system to crystallise into a mixture of polymorphs. The nucleation of ice has been extensively studied in computer simulations. Moore and Molinero,\textsuperscript{10} Reinhardt and Doye,\textsuperscript{11} Li et al.\textsuperscript{12} and Quigley\textsuperscript{13} all found that water crystallised into a mixture of hexagonal ice (the equilibrium and common ice polymorph) and cubic ice (a metastable ice polymorph of only slightly higher free energy). What looks like defected mixtures of cubic and hexagonal ice is seen in experiments where the ice forms at very low temperatures.\textsuperscript{14,15} Malkin et al.\textsuperscript{14} calculated $S(k)$’s for the ice that was a mixture of hexagonal and cubic ice, and found an $S(k)$ that was identifiably crystalline. This is different from our amorphous state which is also a mixture of polymorphs but where the $S(k)$ is liquid-like. At the values of $k$ at which both hexagonal and cubic ice have Bragg peaks, they observed Bragg peaks. It is possible that our finding of a liquid-like $S(k)$ is due to the fact that there are no $k$ values where all three polymorphs have Bragg peaks.
Table S1: Results for attempts to crystallise nine independent realisations of the amorphous state. For each realisation, 200 independent computer simulation runs were started from the amorphous configuration. Configurations are numbered 1 to 10, with no configuration 2. The numbers in columns 3, 4 and 5 give the number of simulation runs that result in a pure fcc crystal, a pure bcc crystal, and no crystallisation, respectively. Column 9 gives the median nucleation time for each configuration. Note that this is not defined for number 7 as less than half the runs resulted in crystallisation (the $P(t)$ curve is in Figure 5). A pure bcc crystal is when $> 80\%$ of the molecules are identified as being bcc. A pure fcc crystal is when $> 80\%$ of the molecules are identified as either fcc or hcp. fcc crystals can form with stacking faults and then can contain significant fractions of molecules identified as hcp. Thus our fcc crystals are predominantly fcc but in some cases have $20\%$ or more hcp molecules due to stacking faults, and this definition allows us to count them together with fcc crystals without stacking faults. Stacking faults are common in the fcc polymorph in the Gaussian Core Model, as they are in other systems of spherically symmetric particles that form fcc crystals, for example hard spheres.\textsuperscript{16,17} The nucleation time is defined to be the first time the 80% threshold is exceeded in a run. The numbers in columns 6, 7 and 8 are the average values of the LD order parameters $\bar{q}_6$, $\bar{w}_6$ and $\bar{w}_4$ in the configuration. All configurations were of 5000 molecules.

<table>
<thead>
<tr>
<th>Config. number</th>
<th>label</th>
<th>fcc</th>
<th>bcc</th>
<th>neither</th>
<th>$\bar{q}_6$</th>
<th>$\bar{w}_6$</th>
<th>$\bar{w}_4$</th>
<th>median nucleation time (MC cycles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pos54-11</td>
<td>200</td>
<td>0</td>
<td>0</td>
<td>0.413</td>
<td>$-0.00466$</td>
<td>$-0.0728$</td>
<td>13,400</td>
</tr>
<tr>
<td>3</td>
<td>pos54-100</td>
<td>200</td>
<td>0</td>
<td>0</td>
<td>0.422</td>
<td>0.000874</td>
<td>$-0.0578$</td>
<td>16,370</td>
</tr>
<tr>
<td>4</td>
<td>pos54-101</td>
<td>117</td>
<td>0</td>
<td>83</td>
<td>0.409</td>
<td>$-0.00389$</td>
<td>$-0.0578$</td>
<td>495,560</td>
</tr>
<tr>
<td>5</td>
<td>pos54-102</td>
<td>102</td>
<td>77</td>
<td>21</td>
<td>0.419</td>
<td>$-0.00131$</td>
<td>$-0.0446$</td>
<td>249,500</td>
</tr>
<tr>
<td>6</td>
<td>pos54-105</td>
<td>0</td>
<td>200</td>
<td>0</td>
<td>0.430</td>
<td>0.00606</td>
<td>$-0.0685$</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>pos54-112</td>
<td>3</td>
<td>86</td>
<td>111</td>
<td>0.411</td>
<td>0.00353</td>
<td>$-0.0230$</td>
<td>not defined</td>
</tr>
<tr>
<td>8</td>
<td>pos54-113</td>
<td>0</td>
<td>200</td>
<td>0</td>
<td>0.417</td>
<td>0.00589</td>
<td>$-0.0613$</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>pos54-114</td>
<td>0</td>
<td>200</td>
<td>0</td>
<td>0.406</td>
<td>0.00359</td>
<td>$-0.0423$</td>
<td>29,400</td>
</tr>
<tr>
<td>10</td>
<td>pos54-117</td>
<td>200</td>
<td>0</td>
<td>0</td>
<td>0.446</td>
<td>$-0.00957$</td>
<td>$-0.0694$</td>
<td>0</td>
</tr>
</tbody>
</table>

**Crystallisation of the amorphous state**

To determine the role of the amorphous state in crystallisation, we produced nine independent realisations of the amorphous state, and attempted to crystallise each one. For each configuration, we ran 200 independent simulation runs, each was of up to $3 \times 10^6$ MC cycles. In some of these runs a pure fcc crystal formed, in others a pure bcc crystal formed. However in some runs neither the fcc nor the bcc formed, and the configuration remained in the amorphous state.

**Most realisations of the amorphous state always crystallise into the same polymorph**

Table S1 details the breakdown for the nine configurations studied. Configurations 1, 3, 4 and 10 formed only the fcc polymorph. Configurations 6, 8 and 9 only formed bcc crystals. Thus in seven
of the nine configurations, the polymorph that forms is (almost) already determined, although which polymorph is selected varies from one realisation of this non-equilibrium state to another.

The runs on one of the remaining configurations, number 5, are roughly evenly split between forming fcc, bcc, or remaining amorphous for whole run, while configuration 7 predominantly forms the bcc polymorph but does occasionally form the fcc polymorph. Thus, in summary, most realisations of the amorphous state are committed to either the fcc or the bcc polymorphs, but some are not. There is also a huge variation in how fast the realisations crystallise, as shown in the median nucleation times in Table S1.

What feature of a realisation of the amorphous state determines which polymorph it will form?

As we noted in the main text, we looked for features of the configuration of molecules in an amorphous state that determines whether it will crystallise into the fcc or bcc polymorphs. We did not find anything that always predicted which polymorph would form. The closest we came to predictive features were the mean value of the LD order parameter \( \bar{w}_6 \) in the configuration, and the numbers of molecules in fcc-like and bcc-like environments. \( \bar{w}_6 \) is the order parameter we use to distinguish between the close-packed and bcc crystals.

The mean values of the order parameters for the nine configurations are shown in Table S1. An obvious question to ask is: Are the average values of \( \bar{w}_6 \) significantly different in the configurations that go on to crystallise into the fcc and bcc polymorphs? Note that, as we see from Figure S1, \( \bar{w}_6 \approx -0.013 \) in well-ordered fcc crystals and 0.013 in the well-ordered bcc crystal. In the nine amorphous configurations \( \bar{w}_6 \) is about an order of magnitude smaller than in either of the well-ordered crystal phases, i.e., its average value is close to zero, the threshold we use to classify molecules as close-packed or bcc.

For the four configurations (1, 3, 4 and 10) that go on to crystallise into the fcc polymorph, the average value of \( \bar{w}_6 = -0.0043 \pm 0.0043 \), where we also show the standard deviation of \( \bar{w}_6 \) in the four configurations. For the three configurations (6, 8 and 9) that go onto to crystallise into the bcc polymorph, the average value of \( \bar{w}_6 = 0.0052 \pm 0.0014 \). So \( \bar{w}_6 \) is slightly negative on average for those configurations that go on to form fcc crystals and slightly positive for those that crystallise into the bcc polymorph, but the values are small, and the value of \( \bar{w}_6 \) is not a perfect predictor: configuration 3 has a positive value of \( \bar{w}_6 \) but forms the fcc polymorph. On average, configurations that go on to crystallise into the fcc (bcc) polymorph do have more fcc-like (bcc-like) ordering, according to our LD order parameters. But this excess ordering is at limit of being detectable, even with our knowledge of the positions of all molecules.

The situation is similar if we look at the numbers of molecules in each polymorph. Table S2 has these numbers. Those configurations that crystallise into the bcc polymorph have more molecules in bcc-like environments than in fcc-like environments. Three of the four configurations that crystallise into fcc, also have more fcc than bcc molecules, but configuration number 3 has more bcc than fcc and hcp molecules. So, again the trend is what you would expect, but the relative numbers of molecules in bcc and fcc environments does not always predict which polymorph forms.

For most polymorphs in most configurations, a large majority of the molecules identified as a particular polymorph, are in a single cluster. So in most configurations there is a single fcc crystal with most of the fcc molecules, a single hcp crystal with most of the hcp molecules, and a single...
Table S2: Results for attempts to crystallise nine independent realisations of the amorphous state. For further details, see the caption to Table S1. Column 2 indicates what the realisation of the amorphous state does: crystallise into the fcc crystal, into the bcc crystal or remain amorphous. For more details see Table S1. Columns 3, 4 and 5 give the total numbers of molecules in fcc, hcp and bcc environments, $T_{fcc}$, $T_{hcp}$ and $T_{bcc}$, respectively. Columns 6, 7 and 8 give the numbers of molecules in the largest fcc, hcp and bcc clusters, $C_{fcc}$, $C_{hcp}$ and $C_{bcc}$, respectively. Note that for most polymorphs in most configurations, the number in the largest cluster are is not much less than the total number. Most of the molecules of a particular polymorph are in a single cluster.

<table>
<thead>
<tr>
<th>Configuration number</th>
<th>xtal</th>
<th>$T_{fcc}$</th>
<th>$T_{hcp}$</th>
<th>$T_{bcc}$</th>
<th>$C_{fcc}$</th>
<th>$C_{hcp}$</th>
<th>$C_{bcc}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>fcc</td>
<td>3032</td>
<td>606</td>
<td>1219</td>
<td>3025</td>
<td>516</td>
<td>1167</td>
</tr>
<tr>
<td>3</td>
<td>fcc</td>
<td>1651</td>
<td>608</td>
<td>2709</td>
<td>1600</td>
<td>547</td>
<td>2708</td>
</tr>
<tr>
<td>4</td>
<td>fcc</td>
<td>2628</td>
<td>812</td>
<td>1320</td>
<td>2612</td>
<td>717</td>
<td>1304</td>
</tr>
<tr>
<td>5</td>
<td>fcc or bcc</td>
<td>1775</td>
<td>1078</td>
<td>2110</td>
<td>1753</td>
<td>1059</td>
<td>2107</td>
</tr>
<tr>
<td>6</td>
<td>bcc</td>
<td>525</td>
<td>153</td>
<td>4320</td>
<td>433</td>
<td>82</td>
<td>4320</td>
</tr>
<tr>
<td>7</td>
<td>amorph.</td>
<td>648</td>
<td>939</td>
<td>3381</td>
<td>532</td>
<td>905</td>
<td>3381</td>
</tr>
<tr>
<td>8</td>
<td>bcc</td>
<td>772</td>
<td>227</td>
<td>3950</td>
<td>710</td>
<td>173</td>
<td>3950</td>
</tr>
<tr>
<td>9</td>
<td>fcc or bcc</td>
<td>1078</td>
<td>491</td>
<td>3334</td>
<td>1016</td>
<td>155</td>
<td>3334</td>
</tr>
<tr>
<td>10</td>
<td>fcc</td>
<td>3416</td>
<td>1316</td>
<td>261</td>
<td>3413</td>
<td>1316</td>
<td>128</td>
</tr>
</tbody>
</table>

The larger clusters often span the periodic boundary conditions of the simulation box, as we can see on the right of Figure S3. When there are two large system-spanning clusters, they form a pair of bicontinuous networks, in an way analogous to that found in bicontinuous microemulsion phases formed in block copolymer and surfactant systems.\textsuperscript{18,19}

**Growth in lengthscale of crystalline ordering during the second step**

It is possible to estimate the size of small crystallites from their $S(k)$. In 1918 Scherrer showed how to estimate the sizes of the crystallites in a polycrystalline sample from the width of the first peak in $S(k)$.\textsuperscript{20} The estimate is based on the location of the first peak, at $k = k_1$, and on the Full-Width-at-Half-Maximum (FWHM) of this peak, $\Delta k_1$. Scherrer showed that the linear size of the crystallites, in units of the lattice spacing, is approximately

\[ \lambda_X \approx k_1 / \Delta k_1 \]  \hspace{1cm} (8)

This method is widely used in experiment as a way of estimating average crystallite size when the samples are polycrystalline and crystallite sizes are on the nanoscale.\textsuperscript{20–22} It is approximate, as the peak width and height are also affected by other factors, such as strain distorting the lattice.

We obtain an estimate of $k_1$ by using the $k$ value where $S(k)$ is a maximum. We then start at $k_1$ and increase $k$ until the first allowed (by box dimensions) value of $k$ at which $S(k) < (1/2)S(k_1)$, this becomes the upper value for the FWHM. We then start at $k_1$ and decrease $k$ in the same way to find the lower value for FWHM. The difference between these upper and lower values is $\Delta k_1$. 

11
Figure S4: Plot of the Scherrer equation estimate for the linear size of crystallites, as a function of time during crystallisation from the amorphous state. The y axis is the ratio of first peak position, $k_1$, to the FWHM of the peak, $\Delta k_1$. The x axis is time in MC cycles. This is for configuration 1 with 5000 molecules. The first peak in $S(k)$ is at $k_1 = 3.06/\sigma$.

At our relatively small system sizes (5000 molecules, about 40 $\sigma$), our $S(k)$’s are noisy due to statistical fluctuations. This noise tends to cause us to underestimate $\Delta k_1$, and so overestimate the effective size of the crystallites, $n$. The noise is reduced by the fact that we use $S(k)$’s smoothed using the window filter described in the previous section.

In Figure S4, we have plotted $k_1/\Delta k_1$ as a function of time during crystallisation from the amorphous state. We see that in the initial amorphous state (at time = 0) we estimate that the effective size of the crystallites is of order ten lattice spacings across. During crystallisation, this size increases relatively slowly until towards the end when the simulation box is almost entirely fcc, when the effective size rapidly increases. The size then saturates as our calculated $\Delta k_1$ cannot be smaller than the spacing between allowed values $k$ values in our finite box (see section above on how we calculate $S(k)$) — i.e., this saturation is a finite-size effect.

In summary, although there is no unique way to measure the size of small crystalline domains, the Scherrer equation, Eq. (8) gives an estimated value of around ten molecules across in the amorphous state. Also, when the system has crystallised, it gives a value larger than our system size of around 40 $\sigma$.

**Timescales for growth of the amorphous state and crystallisation of the amorphous state**

It takes an average of 53,900 MC cycles (standard deviation of 23,900 cycles, over nine runs) for the amorphous state to grow from just past the nucleation barrier, to completely fill our simulation box. This is in a system of 5000 molecules. Here we define just past the nucleation barrier to be a largest cluster of crystalline molecules of 540 molecules; top of the barrier is approximately at a cluster of 380 molecules. We define complete conversion as 90% of the molecules in the system crystalline.
The median nucleation time for our nine configurations is quite short, 18,140 MC cycles (N.B. as not all 1800 runs crystallise the mean nucleation time is not defined). As this is less than the average time to completely fill the box, this means that in a majority of cases, the amorphous state is transient. However, there is huge variability from one configuration to another. The median nucleation times for each of the nine configurations are shown in Table S1.

The rate at which a system of 5000 molecules in the amorphous state crystallises varies very widely, both from one of the nine configurations to another, and from one run on a given configuration, to another run on the same configuration. Some runs always crystallised essentially immediately (configurations 6 and 10) while the median nucleation time for configuration 4 is almost 500,000 cycles and that for configuration 7 cannot be defined as only 45% of the runs resulted in crystallisation. Median nucleation times are in Table S1.

Configuration 5 crystallised into both polymorphs large numbers of times, into fcc 107 times and bcc 77 times. Thus for this configuration (only) we could compare the timescales for crystallisation into the two competing polymorphs. The timescales are very different, the fcc polymorph formed much faster than the bcc polymorph. The mean crystallisation times are 173,000 MC cycles for fcc, and 628,000 for bcc. The standard deviations are 109,000 MC cycles for fcc, and 460,000 for bcc. So although in both cases there are large spreads in the nucleation times, the means are clearly different. The estimated uncertainties in the means are \( \approx \frac{109,000}{\sqrt{107}} = 11,000 \) for fcc and \( \approx \frac{460,000}{\sqrt{77}} = 52,000 \) for bcc, and so these uncertainties do not overlap. We do not know why the timescales for crystallisation are so different.

**Estimation of survival function \( P(t) \)**

In Figure 5 of the main text we plot the survival function, \( P(t) \), of the amorphous state, both for all nine configurations combined and for individual configurations. We estimate \( P(t) \) using the Kaplan-Meier estimator, which is standard in the analysis of survival data.\(^{23,24}\) For a set of \( n_{\text{RUN}} \) runs (200 for an individual configuration, 1800 for all nine combined), we have a set of \( n_{\text{RUN}} \) nucleation times, \( t_i \), for \( i = 1, n_{\text{RUN}} \). We assume that these times are ordered, with \( i = 1 \) the shortest nucleation time, and \( i = n_{\text{RUN}} \) the last one. \( t_i \) for runs are taken as at infinity if the run never crystallises. We also define \( n_i \) as being the number of runs that have not crystallised immediately before nucleation time \( t_i \), for example for an individual configuration, \( n_i = 200 - (i - 1) \). Then the Kaplan-Meier estimator for \( P(t) \) is defined by\(^{23,24}\)

\[
P(t) = \prod_{t_i < t} \frac{n_i - 1}{n_i}
\]

(9)

where the product at a time \( t \) is over all nucleation times \( t_i \) shorter than \( t \).

**Larger system of 50,000 molecules**

Although most of our results are for relatively small system sizes of 5000 molecules, we do have a few studies of the amorphous state in larger systems of 50,000 molecules. A snapshot and structure factor for a system of this size is shown in Figure S5. As for our smaller systems, the state is a mixture of all three polymorphs, and has a liquid-like \( S(k) \).

In Figure S5, we have also plotted two-dimensional structure factors in all three Cartesian
Figure S5: Snapshot, structure factors and pair distribution function for a large, 50,000 molecules, system. (a) snapshot of the simulation. (b) is a powder $S(k)$ while (c), (d) and (e) are two-dimensional structure factors, $S(k_x, k_y, 0)$, $S(k_x, 0, k_z)$ and $S(0, k_y, k_z)$, respectively. Here $k = (k_x, k_y, k_z)$. Note that the system is at the state point $T = 0.0262\epsilon/k_B, P = 0.01\epsilon/\sigma^3$. This is different from the state point studied in the rest of this work, which is $T = 0.002\epsilon/k_B, P = 0.0055\epsilon/\sigma^3$. Composition is 37% fcc, 12% hcp, 42% bcc and 9% liquid. In (b) the smoothed $S(k)$ is plotted as a black curve. For this larger system size, smoothing has relatively little effect, except for eliminating some outliers, so we have also plotted, as red circles, the raw unsmoothed $S(k)$ for comparison. (f) The pair distribution function $g(r)$ for the configuration. This configuration is available in .xyz format as part of the supplementary material.
planes: $S(k_x, k_y, 0)$, $S(k_x, 0, k_z)$ and $S(0, k_y, k_z)$. We have smoothed them by averaging over squares with sides of length seven allowed (by the box length) components.

Note that the structure factors are not isotropic. All three of Figure S5(c), (d) and (e) have two peaks in the first ring at $k \approx 3.5$, and this ring in (d) in particular has some structure with six-fold symmetry. Systems of this small size, approximately $80\sigma$ across, do not contain enough crystal domains for the structure factor to average out into the rings characteristic of polycrystalline samples. Thus, if in an experiment diffraction patterns could be obtained from sufficiently small volumes, the structure factor would not be isotropic. It could then be distinguished from the more isotropic structure factor of a liquid, or an amorphous state with truly liquid-like structure. Diffraction patterns obtained from volumes containing many randomly oriented nanocrystalline domains will be isotropic, and so only Figure S5(b) will be obtained, not the angular information seen in Figure S5(c), (d) and (e). Volumes containing many domains will give $S(k_x, k_y, 0)$, $S(k_x, 0, k_z)$ and $S(0, k_y, k_z)$ functions that contain only rings, not spots.

Figure S5(f) is the pair distribution function for the configuration. This is a decaying oscillatory function, as it is in the liquid state, but the decay is slower (consistent with the greater ordering predicted by the higher first peak in $S(k)$); note that even up to $r = 30\sigma$, approximately 15 times the nearest-neighbour distance, there are still detectable although very weak oscillations.

**Videos**

Files video_config5.mp4 is a zoomed in animation for the configuration in Figure 1 of the main paper.

**Two configurations in .xyz format**

Configuration number 5 (5,000 molecules) and the large configuration shown in Figure S5 (50,000) molecules are both available in .xyz format as part of the supplementary information. They are config5.xyz and largeconfig1.xyz, respectively. In these files, the liquid, fcc, hcp and bcc molecules are given as ‘N’, ‘S’, ‘P’ and ‘F’, respectively.

**References**


