Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics


J. Phys. Chem. Lett., Just Accepted Manuscript • DOI: 10.1021/acs.jpclett.8b00902 • Publication Date (Web): 12 May 2018

Downloaded from http://pubs.acs.org on May 14, 2018

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.
Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics

Volker L. Deringer,†,‡,* Noam Bernstein,† Albert P. Bartók,§ Matthew J. Cliffe,†
Rachel N. Kerber,‡ Lauren E. Marbella,‡ Clare P. Grey,‡ Stephen R. Elliott,‡
and Gábor Csányi†

† Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom
‡ Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom
¶ Center for Materials Physics and Technology, U.S. Naval Research Laboratory,
   Washington, DC 20375, United States
§ Scientific Computing Department, Science and Technology Facilities Council,
   Rutherford Appleton Laboratory, Oxfordshire OX11 0QX, United Kingdom

*Corresponding author. E-mail: vld24@cam.ac.uk
ABSTRACT: Amorphous silicon (a-Si) is a widely studied non-crystalline material, and yet the subtle details of its atomistic structure are still unclear. Here, we show that accurate structural models of a-Si can be obtained using a machine-learning-based interatomic potential. Our best a-Si network is obtained by simulated cooling from the melt at a rate of $10^{11}$ K/s (that is, on the 10 ns timescale), contains less than 2% defects, and agrees with experiments regarding excess energies, diffraction data, and $^{29}$Si NMR chemical shifts. We show that this level of quality is impossible to achieve with faster quench simulations. We then generate a 4,096-atom system which correctly reproduces the magnitude of the first sharp diffraction peak (FSDP) in the structure factor, achieving the closest agreement with experiments to date. Our study demonstrates the broader impact of machine-learning potentials for elucidating structures and properties of technologically important amorphous materials.
Amorphous silicon (a-Si) is a fundamental and widely studied non-crystalline material, with applications ranging from photovoltaics and thin-film transistors to electrodes in batteries.\textsuperscript{1–5} Its atomic-scale structure is traditionally approximated in a Zachariasen-like picture\textsuperscript{6} with all atoms in locally “crystal-like”, tetrahedral environments, but without long-range order.\textsuperscript{7–9} However, the real material contains a non-zero amount of coordination defects, colloquially referred to as “dangling bonds” (under-coordinated sites) and “floating bonds” (over-coordinated sites). Knowing the properties and abundance of such defects is important, as they can control electronic and other macroscopic properties. We note at the outset that, although defect sites in a-Si may be passivated by hydrogenation (to give “a-Si:H”) in some synthetic conditions, we here focus on the archetypical, hydrogen-free material as made in ion-implantation or sputter-deposition experiments.\textsuperscript{10–14}

Even the most advanced experimental approaches do not directly allow the observation of the bulk atomic structure in amorphous materials. Despite significant advances, including \textit{in situ} NMR techniques\textsuperscript{15,16} and “inverse” approaches such as Reverse Monte Carlo (RMC) modeling of diffraction data,\textsuperscript{17–19} only indirect knowledge can be gained about the local atomic environments, and that only in a statistical sense. For almost three decades, molecular-dynamics (MD) simulations have therefore played a crucial and complementary role, with a-Si being a prominent example.\textsuperscript{20–24} These simulations either use density-functional theory (DFT) or classical force fields. DFT-MD describes a system with quantum-mechanical accuracy and can largely correctly capture the structural and bonding subtleties of liquid and amorphous matter. However, it is computationally expensive, and therefore allows only limited system sizes (a few hundred atoms at most) and timescales to be simulated. Indeed, the cooling rates in previous DFT simulations of a-Si ($\approx 10^{14}$ K/s) are orders of magnitude faster than those in experiments.\textsuperscript{20–22} Classical force fields require much less computational effort, giving access to nanometer-scale (“device-size”) struc-
tural models,\textsuperscript{9} both for $a$-Si and for multicomponent systems derived from it (see ref 25 for but one example). However, they are rarely accurate enough to fully describe the structural variations present in the amorphous state.

Capitalizing on today’s “big-data” revolution, machine-learning (ML) algorithms are increasingly used to generate interatomic potentials for atomistic simulations.\textsuperscript{26–30} By “learning from” (or rather, \textit{fitting to}) quantum-mechanically computed reference data for energies and forces, ML-based interatomic potentials can enable simulations with an accuracy that is largely comparable to DFT, but with a computational cost that is orders of magnitude lower, and with linear (order-$N$) scaling behavior. Comparison to experimental observables is thereby the ultimate benchmark and means of validation for the quality of any ML-based interatomic potential, as we stress that no experimental but only DFT-computed data enter the “learning” process.

We believe that such ML potentials are particularly promising for disordered and amorphous materials, which must be represented by nanometer-scale structural models containing several hundreds or thousands of atoms. A landmark example has been the development of an artificial neural-network potential for the phase-change material GeTe,\textsuperscript{31} enabling simulation of the crystallization properties\textsuperscript{32} including entire nanowires.\textsuperscript{33} We recently introduced an ML potential for amorphous carbon,\textsuperscript{34} based on the Gaussian approximation potential (GAP) framework\textsuperscript{27} and the Smooth Overlap of Atomic Positions (SOAP) atomic similarity kernel,\textsuperscript{35} which captures the intricate structural, mechanical, and surface properties of the material\textsuperscript{34} and, more recently, has enabled accurate large-scale simulations of the growth mechanism.\textsuperscript{36} Very recent work using neural-network potentials allowed for the atomistic modeling of amorphous Li$_x$Si phases relevant in battery applications.\textsuperscript{37,38} Finally, such potentials were used in seminal studies to describe the complex phase transitions between polymorphs of \textit{crystalline} Si.\textsuperscript{26,39}
In this Letter, we show how realistic atomistic modeling of a-Si can be enabled by an ML-based interatomic potential, again using SOAP and GAP. We first report on melt–quench simulations with cooling rates much slower (that is, better) than what can be achieved in quantum-mechanical-based simulations, and we show how this leads to a higher-quality and lower-energy structure of a-Si. Our structural models show excellent agreement with experiments probing local structure, including $^{29}$Si NMR shifts and diffraction data for high-quality samples, and open the door for future combined modeling and experimental studies on disordered and amorphous materials.

Simulated quenching from the melt is a widely used technique for generating amorphous model networks. In this, one starts with a liquid and progressively lowers the temperature, “freezing in” an amorphous structure. However, for silicon, this approach is not trivial, due to the change in local environments between the high-coordination metallic liquid and the tetrahedral-like amorphous state. We decided to perform a set of variable-volume and constant-pressure (NPT) quench simulations, in which we varied the quench rate, and thus the run-time, by several orders of magnitude. These were carried out using LAMMPS;\textsuperscript{40} details are in the Supporting Information. For the moment, we focus on a system size of 512 atoms in the cell and perform a single simulation at each quench rate. This system size is significantly larger than what has so far been accessible to DFT (64–216 atoms),\textsuperscript{20–22,24} but smaller than what is possible for empirical potentials; this will be addressed directly later on. Our fastest quench rate ($10^{14}$ K/s) corresponds to early, seminal DFT studies,\textsuperscript{20–22} whereas our simulations at $10^{12}$ K/s mirror the limit of what is presently possible for DFT-quality MD. In contrast, we here use a recently developed GAP model\textsuperscript{30,41} which allows us to increase the simulation time tenfold beyond that, namely decreasing the quench rate to $10^{11}$ K/s, while retaining similar accuracy.
While an increase in simulation time by one order of magnitude may seem incremental at first sight, the full power of ML potentials becomes apparent when looking at the overall computational effort required (Figure 1a). For demonstration, we performed a brief DFT-MD simulation on a 512-atom a-Si network and use the timing information for a rough extrapolation (Supporting Information). Quenching with a rate of $10^{11}$ K/s would thus require around 16 million core hours, or current nominal costs of $185,000 on the UK national supercomputer. In contrast, the same quench rate in GAP-MD required below 40,000 core hours, equivalent to nominal costs below $500. Using GAP, it would hence be possible to decrease the quench rate even further, but given the results obtained at $10^{11}$ K/s we subsequently chose to increase the system size instead (see below).

**Figure 1.** Unlocking slow quenching in molecular-dynamics simulations, using the GAP machine-learning framework, and its application to a-Si. (a) Computing time required for MD simulations with different quench rates. Decreasing the quench rate by an order of magnitude increases the number of required MD steps, and thus the CPU cost, by the same factor. Quench rates of $\approx 10^{12}$ K/s have so far been the limit for 512-atom DFT-MD simulations, and a system size of
4,096 atoms ("4k") has been widely out of reach. Both limits can be overcome using GAP. Our "4k" simulation (magenta) employs an adapted temperature protocol; DFT timing information is extrapolated from a short trajectory; see Supporting Information. (b) Stability of 512-atom $a$-Si structures, taken at various stages of GAP melt–quench trajectories and subsequently relaxed; energies given relative to crystalline (diamond-type) Si. Experimental data refer to samples freshly deposited ("as-dep.") or annealed at progressively higher temperatures.\textsuperscript{12} (c) Angle distribution functions for $a$-Si GAP structures. Points show original data, sampled from short (5 ps) MD simulations; lines show Gaussian fits; data for different quench rates are vertically offset for clarity. (d) Medium-range order in these $a$-Si networks, assessed by shortest-path ring statistics.\textsuperscript{42}

The slow quench rate of $10^{11}$ K/s, “unlocked” here using GAP, is indeed required to generate reliable structural models of $a$-Si. This is seen in Figure 1b: we took structural snapshots at various increments of the simulations, optimized them into local minima, and plotted their energy (relative to the thermodynamically stable form, diamond-type $c$-Si) as a function of how far the quench has progressed in time from the liquid to the final $a$-Si structure. The right-hand side shows the experimental sample stability with increasing annealing and thus ordering, based on calorimetry (as is common, we approximate $\Delta E \approx \Delta H$ when comparing theory and experiment). Intermediate quench rates lead to $a$-Si networks that are as stable as freshly deposited or partially annealed samples ($\Delta E \approx 0.17–0.20$ eV/atom). In contrast, our slowest quench at $10^{11}$ K/s yields a structure whose stability matches the experimental result for a well-annealed sample from ref 12 ($\Delta E \approx 0.14$ eV/atom). The GAP-computed bulk moduli for these $a$-Si networks range from 62–69 GPa and increase with slower quenching (the material becoming “harder”); the computed Young’s moduli increase from 73 to 98 GPa; see Supporting Information.

The benefit of slow quenching is further seen in two of the most common structural indicators used for amorphous solids. In $a$-Si, the bond angles are distributed around the ideal tetrahedral value (109.5°; Figure 1c). Fitting Gaussian distributions to these data allows us to determine the full width at half maximum (FWHM), which decreases gradually from 30° to 22° with in-
creasingly slower quenching. The experimental value for the bond-angle deviation of \( \approx 11^\circ \) (ref 10) is consistent with the half width at half maximum (HWHM) for our slowest quench. Moreover, the medium-range structural order is important in covalent amorphous networks,\textsuperscript{43} and we quantify it here using shortest-path ring statistics.\textsuperscript{42} In diamond-type \( c\)-Si, all atoms are in six-membered ring (cyclohexane-like; \( m = 6 \)) configurations, whereas \( a\)-Si also contains a large number of five- and seven-membered rings, and a lesser amount of smaller and larger ones. All rings with \( m \neq 6 \) depart from the reference crystalline state, and as such are a measure of disorder, but we here distinguish them further as follows. Five- and seven-membered rings are still expected to be energetically viable (supported by their abundance in \( a\)-Si), whereas, for example, four-membered rings will be clearly under strain. We therefore label rings with \( m < 5 \) as “small-ring defects”, and rings with \( m > 7 \) as “large-ring defects” (Figure 1d).

**Figure 2.** Progressive ordering in \( a\)-Si, comparing 512-atom GAP structures to experiments. (a) Computed structure factor \( S(Q) \) (purple) and X-ray diffraction data for a well-annealed sample (gray; digitized from ref 14). (b) Close-up around the third peak, in which data for the different quench rates have been offset vertically and are each compared to the same experimental dataset.
(points).\textsuperscript{14} (c) Solid-state $^{29}$Si NMR chemical shifts, computed for GAP-quenched structures (raw data as histograms; Gaussian fits as colored lines), and experimental data from ref 44 (bottom).

In Figures 2a–b, we show computed structure factors, $S(Q)$, which can be compared to diffraction experiments. The third peak (at $\approx 5–7$ Å$^{-1}$) gradually splits into two well-defined sub-peaks when moving from the $10^{14}$ K/s (yellow) to the $10^{11}$ K/s data (purple). This is qualitatively consistent with experimental observations: as-deposited samples show a fairly featureless third peak, whereas annealed ones (and also our $10^{11}$ K/s result) exhibit a clear splitting into sub-peaks.\textsuperscript{14} Even better agreement with the experimental structure factor can be achieved for a larger structural model containing 4,096 atoms, which we will show below.

We furthermore computed solid-state $^{29}$Si NMR chemical shifts, $\delta$, for all atoms in the unit cells, thereby characterizing each atomic environment individually. We use established DFT-based algorithms\textsuperscript{45–47} and reference all $\delta$ values to tetramethylsilane (TMS), analogous to experiments. The results for the different GAP structures are shown in Figure 2c (histograms). Furthermore, due to the broad distribution of $\delta$ values in the amorphous state, we fit Gaussian profile functions to these data (lines), as detailed in the Supporting Information. We compare the output of these computations to experiments for pure $\alpha$-Si prepared by sputter deposition.\textsuperscript{44} The latter samples were analyzed via secondary-ion mass spectrometry (SIMS), showing no measurable oxygen contamination and $\approx 0.2$ at.-% hydrogen in the samples.\textsuperscript{44} This low level of impurity is thought to have little or no impact on the $^{29}$Si NMR results, enabling direct comparison to our simulations. In addition to the numerical values reported in ref \textsuperscript{44}, we fit a Gaussian profile to the experimental data for the sample annealed at 520 °C (before the onset of crystallization at higher temperature). We perform this fit using the same procedure as for our DFT data (Table 1). This yields numerical quality criteria that can be used to assess any given structural model.
Clearly, simulations using the two fastest quench rates (yellow and orange) lead to structures with very large scatter in the computed NMR shifts, as a direct consequence of their distorted atomic environments (and thus large FWHM for the Gaussian fits). The network generated at a slower quench rate, 10^{12} K/s (red), agrees more appreciably with experiment with regard to both the broadness and the center of mass for the Gaussian fit (\(\delta_{\text{DFT}} = -37 \text{ ppm}\)); the latter can be compared to \(\delta_{\text{exp}} = -38.3 \text{ ppm}\) for as-deposited \(\alpha\)-Si, and \(\delta_{\text{exp}} = -42.9 \text{ ppm}\) for a sample annealed at 580 °C.\(^{44}\) Hence, there is a progressive shift to lower frequency in the experimental data with increasing structural ordering, and this is reproduced by our quenched structure at 10^{11} K/s (\(\delta_{\text{DFT}} = -51 \text{ ppm}\)), both qualitatively and quantitatively (to within a few ppm). The results for the 10^{11} K/s quench also compare well with those for a DFT-optimized Wooten–Winer–Weaire (WWW) network of \(\alpha\)-Si (\(\delta_{\text{DFT}} = -53 \text{ ppm}\); structure taken from ref 48), while those for the faster quenches do not (Table 1).

<table>
<thead>
<tr>
<th>Table 1.</th>
<th>(^{29})Si NMR Parameters of (\alpha)-Si from Simulation and Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\delta) (ppm)</td>
<td>Gaussian FWHM (ppm)</td>
</tr>
<tr>
<td>GAP-MD, 10^{14} K/s quench</td>
<td>-20</td>
</tr>
<tr>
<td>GAP-MD, 10^{13} K/s quench</td>
<td>-38</td>
</tr>
<tr>
<td>GAP-MD, 10^{12} K/s quench</td>
<td>-37</td>
</tr>
<tr>
<td>GAP-MD, 10^{11} K/s quench</td>
<td>-51</td>
</tr>
<tr>
<td>Expt. (as-dep.; ref(^{44}))</td>
<td>-38.3</td>
</tr>
<tr>
<td>Expt. (580 °C annealed; ref(^{44}))</td>
<td>-42.9</td>
</tr>
<tr>
<td>Expt. (520 °C annealed; Gaussian fit)</td>
<td>-45</td>
</tr>
<tr>
<td>Relaxed WWW model</td>
<td>-53</td>
</tr>
</tbody>
</table>
We now place our melt–quench simulations into a wider context, as there are several different ways of modeling \( \alpha \)-Si. First, we survey results of RMC modeling, which is an established means of extracting structural information from diffraction data\(^{17}\). Recent work by some of us showed that reasonable restraints can improve the RMC modeling of \( \alpha \)-Si\(^{18}\). In particular, the SOAP similarity measure, initially developed to encode atomic structure in ML potentials\(^{35}\), proved useful for this purpose\(^{48}\). SOAP-RMC output, subsequently relaxed using DFT, has thus been shown to provide a high-quality structural model of \( \alpha \)-Si\(^{48}\). We now take the same structures but anneal them further using GAP: heating to 1,100 K, holding, and cooling back to 300 K, for a total simulation time of 50 ps. This relatively short annealing is thought to be appropriate, as a recent DFT-MD study showed that annealing a quenched structure at 10 ps versus 20 ps had no appreciable effect on the outcome\(^{24}\). We also performed the same annealing procedure for the DFT-optimized WWW model from ref 48; a somewhat similar strategy has been followed before, based on a tight-binding model and a system size of 216 atoms\(^{23}\). Finally, we include a state-of-the-art 216-atom structure that was carefully generated in a recent work, by slow quenching using the empirical Tersoff potential and subsequent multistep optimization using DFT (here labeled “Tsf+DFT”)\(^{24}\).

We compare these structures in Figure 3 using three types of quality indicators. First, we report the number of coordination defects (Figure 3a), counting three- and fivefold bonded atoms with a bond-length cutoff of 2.85 Å. We then measure the distortion from ideal tetrahedral coordination environments: by fitting Gaussians to the angle distributions and determining their FWHM, and by using a numerical order parameter\(^{49}\) that was employed earlier for tetrahedral environments in liquid water\(^{50}\) and chalcogenide glasses\(^{51}\) (Figure 3b). Beyond nearest-neighbor environments, we quantify the medium-range order using shortest-path ring statistics, as above\(^{42}\), again considering as “defects” any rings with fewer than 5 or more than 7 members (Figure 3c).
In all cases, the GAP-quenched structure with the slowest quench rate \((10^{11} \text{ K/s})\) exhibits very good figures of merit. Interestingly, the count of five-membered rings \((m = 5)\) decreases continuously in progressively more ordered GAP-quenched structures, but that of seven-membered rings \((m = 7)\) increases instead, as shown on the far right of Figure 3c.

**Figure 3.** A survey of \(a\)-Si structural models, using common quality criteria; for all of them, the lower, the better. (a) Coordination defects, reporting the count of “dangling bonds” (threelfold bonded atoms) and “floating bonds” (fivelfold bonded atoms). (b) Deviations from tetrahedral geometry, quantified via the FWHM of the fitted angle distribution \((\text{top}; \text{see also Figure 1c})\) and a variant of the Chau–Hardwick order parameter, \(S_{\text{CH}} \,(\text{bottom})\). (c) Count of small- and large-ring defects, as defined in Figure 1d. Smaller plots additionally characterize the number of five- and seven-membered rings in the GAP-quenched structures. In all panels, light gray bars refer to structures from ref \(48\), generated using pure reverse-Monte Carlo (“RMC”), INVERT restraints (“INV”) \(^{18}\) or SOAP restraints. Structures denoted “rel.” have been DFT-relaxed in ref \(48\); structures denoted “ann.” have been further annealed using GAP. “Q” denotes our GAP quenches at the different rates. All structures contain 512 atoms in the cell, except for the Tsf+DFT structure (216 atoms), due to the higher computational cost of DFT annealing.
Finally, we prepared a larger \( a \)-Si structural model containing 4,096 atoms (Figure 4a), using GAP-MD and a variable quench rate between \( 10^{11} \) and \( 10^{13} \) K/s, as detailed in the Supporting Information. This system size is in reach for ML-based interatomic potentials,\(^{32}\) as they are linearly scaling with system size due to their finite cutoff radius (cf. Figure 1a). Having access to \textit{ab initio} quality structural models on the 4-nm length scale allows us to study the medium-range order more closely. This fundamental question has been discussed in recent work on nearly hyper-uniform networks,\(^{52,53}\) in particular, by quantifying the inverse height \( (H^{-1}) \) of the first sharp diffraction peak in the structure factor at around 2 Å\(^{-1}\). This quantity is taken as a measure for the degree of structural ordering.\(^{52,53}\)

We compare our structure with the current state of the art, viz. \( a \)-Si systems containing 100,000 atoms,\(^{9,54}\) in Figure 4b. Although the latter were generated with an improved WWW algorithm, not by slow quenching, they allow us to place our work in the context of existing ultra-large structural models. Surprisingly, the latter system size alone does not seem to be needed if the structural modeling itself is sufficiently accurate. Indeed, looking at \( H^{-1} \), our GAP approach outperforms the previous simulation results in much larger cells, and leads, again, to almost quantitative agreement with experiment \( (H^{-1} = 0.58 \text{ with GAP}, \ H^{-1} = 0.57 \text{ in experiment}; \text{Figure 4b}) \).

By comparison, an \( a \)-Si structure of the same size (4,096 atoms) but generated using empirical potentials gave a much larger \( H^{-1} = 0.81 \) (ref 54). Moreover, our slowest-quenched GAP-based system, even smaller with 512 atoms/cell, yields \( H^{-1} = 0.66 \), remarkably still outperforming the 100,000-atom structure from ref 14 \( (H^{-1} = 0.68) \). Beyond the first sharp diffraction peak alone, Figure 4b also shows that the agreement in the structure factor between the 4,096-atom GAP system and experimental data at larger \( Q \) is excellent, and significantly better than for the VBSB 100,000-atom system.\(^{9}\)
Figure 4. A nanoscale structural model of $\alpha$-Si containing 4,096 atoms, quenched using GAP-MD at a variable rate between $10^{11}$ and $10^{13}$ K/s (Supporting Information). (a) Ball-and-stick drawing, emphasizing the low concentration of coordination defects. The vast majority of atoms are fourfold coordinated ($N = 4$) and are shown as colorless spheres. (b) Reciprocal-space fingerprints in the structure factor, comparing to results for two of the largest structural models to date (containing 100,000 atoms, “100k”) and to experimental data from ref 14. The height of the first sharp diffraction peak, $H$, serves as an indicator for structural ordering; in accord with previous literature, we plot $H^{-1}$ values in the inset. Data for the structure labeled “VBSB” is taken from Vink et al.; data for the structure labeled “HST” is from a more recent study by Hejna, Steinhardt, and Torquato.
In conclusion, we have shown that machine-learning-based interatomic potentials can lead to an unprecedented level of quality in the structural modeling of amorphous materials. We used a Gaussian approximation potential (GAP) to generate high-quality atomistic models of amorphous silicon, quenching from the liquid at a rate of $10^{11}$ K/s, hitherto inaccessible to DFT-quality simulations. These structural models agree convincingly with calorimetry and $^{29}$Si NMR experiments, and with X-ray structure factors, including the height of the first sharp diffraction peak. We note that ML potentials are critically dependent on the quality of the quantum-mechanical input data, and as of today require significant effort to be developed in the first place; in the present case, our GAP has “seen” diverse liquid and amorphous configurations and interpolates between these. These findings will have implications for future research on disordered and amorphous materials, opening the door for quantitatively accurate atomistic modeling with direct links to experiments, for $a$-Si and beyond.

**Acknowledgements**

We thank Professor Jonathan Yates for useful discussions. V.L.D. acknowledges a Feodor Lynen fellowship from the Alexander von Humboldt Foundation, a Leverhulme Early Career Fellowship, and support from the Isaac Newton Trust. N.B. acknowledges support from the Office of Naval Research through the U. S. Naval Research Laboratory’s basic research base program. A.B.P. acknowledges support from the Collaborative Computational Project for NMR Crystallography, funded by EPSRC grant EP/M022501/1. M.J.C. acknowledges support from Sidney Sussex College, University of Cambridge. L.E.M. acknowledges financial support through a FP7 Marie Curie International Incoming Fellowship. Via our membership of the UK’s HEC Materials Chemistry Consortium, which is funded by EPSRC (EP/L000202/1), this work used the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk). **Data access statement:** Original data supporting this publication are provided as Supporting Information.
Associated Content

Supporting Information. Computational details; supplementary data and discussion; coordinate files for structural models.

References


(11) Roorda, S.; Doorn, S.; Sinke, W. C.; Scholte, P. M. L. O.; van Loenen, E. Calorimetric


