Modeling the Phase-Change Memory Material, $Ge_2Sb_2Te_5$, with a Machine-Learned Interatomic Potential

Journal:	*The Journal of Physical Chemistry*
Manuscript ID	jp-2018-06476z.R1
Manuscript Type:	Article
Date Submitted by the Author:	24-Aug-2018

Complete List of Authors:	Mocanu, Felix; University of Cambridge Department of Chemistry; University of Cambridge, Engineering Laboratory Konstantinou, Konstantinos; University of Cambridge Department of Chemistry Lee, Tae Hoon; University of Cambridge Department of Chemistry Bernstein, Noam; Naval Research Laboratory, Deringer, Volker; University of Cambridge, Engineering Laboratory; University of Cambridge Department of Chemistry Csányi, Gábor; University of Cambridge, Engineering Laboratory Elliott, Stephen; University of Cambridge,

Modeling the Phase-Change Memory Material, $Ge_2Sb_2Te_5$, with a Machine-Learned Interatomic Potential

Felix C. Mocanu,[†,‡] *Konstantinos Konstantinou,*[†] *Tae Hoon Lee,*[†] *Noam Bernstein,*[¶] *Volker L. Deringer,*[†,‡] *Gábor Csányi,*[‡,*] *and Stephen R. Elliott*[†,*]

[†] Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom

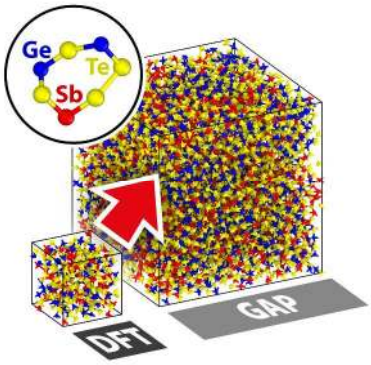
[‡] Engineering Laboratory, University of Cambridge, Cambridge CB2 1PZ, United Kingdom

[¶] Center for Materials Physics and Technology, U.S. Naval Research Laboratory, Washington, DC 20375, United States

*Corresponding authors. E-mail: gc121@cam.ac.uk, sre1@cam.ac.uk

ABSTRACT: The phase-change material, $Ge_2Sb_2Te_5$, is the canonical material ingredient for next-generation storage-class memory devices used in novel computing architectures, but fundamental questions remain regarding its atomic structure and physico-chemical properties. Here, we introduce a machine-learning (ML)-based interatomic potential that enables large-scale atomistic simulations of liquid, amorphous, and crystalline $Ge_2Sb_2Te_5$ with an unprecedented combination of speed and density-functional theory (DFT) level of accuracy. Two applications exemplify the usefulness of such an ML-driven approach: we generate a 7,200-atom structural model, hitherto inaccessible with DFT simulations, that affords new insight into the medium-range structural order; and we create an ensemble of uncorrelated, smaller structures, for studies of their chemical bonding with statistical significance. Our work opens the way for new atomistic insights into the fascinating and chemically complex class of phase-change materials that are used in real non-volatile memory devices.

TOC Graphic



Introduction

The phase-change memory materials of interest in this work are chalcogenide alloys with a fast and reversible transition between a low electrical-conductivity amorphous phase and a metastable, high-conductivity crystalline phase. The Ge-Sb-Te (GST) compound, $Ge_2Sb_2Te_5$, along the GeTe – Sb_2Te_3 pseudo-binary compositional tie-line, has been the canonical phase-change material of this type, first optimized for optical-storage media and more recently used as a basis for non-volatile electronic memory (phase-change random access memory, PCRAM) and consequent data-storage applications.[1–3]

Computer modeling of this material, using first-principles (*ab initio*) molecular-dynamics (MD) simulations, has elucidated atomic details of the liquid, amorphous and crystalline phases, as well as of the amorphous-crystalline phase transition, and provided atomistic insight to guide the design of these materials for different applications.[4–7] This approach, however, is computationally very demanding and suffers from finite-size effects due to the small system sizes that are tractable for DFT simulations. Even the largest DFT-driven MD simulations to date are restricted to model sizes below 1000 atoms in size, as run-times of several nanoseconds are required to observe crystallization in large models.[8] For liquid and amorphous configurations, there is an added challenge of accumulating sufficient statistics from finite-sized models to ensure statistically reproducible results. Overcoming these limitations, without reducing accuracy, has been an outstanding challenge for the wider modeling community.

For the simple, binary prototype phase-change material, GeTe, two types of interatomic potentials have been reported. One is a modified Tersoff-type potential,[9] empirically fitted and applied to study the time-dependent resistance drift in amorphous GeTe using parallel tempering.[10] The other is based on the artificial neural-network approach of Behler and Parrinello,[11] which reaches near-DFT accuracy for the potential by "learning" directly from the quantum-mechanical

data obtained from a large number of configurations, calculated using DFT methods. This neural-network potential by Sosso et al.[12] was initially used to study the atom-dynamical and thermal properties of bulk amorphous GeTe in simulation cells containing up to 8,192 atoms.[13–15] More recently, it was further augmented to describe surfaces and nanostructures.[16] However, neither approach has been able to be extended, so far, to a ternary Ge−Sb−Te composition.

The additional element compared to GeTe, namely Sb, can make a substantial difference to the structure and properties of the material. Despite the fact that $Ge_2Sb_2Te_5$ has been the canonical composition for PCRAM devices, many unanswered questions remain about its structure in the liquid and amorphous phases. Additional structural insight from a statistically significant number of models could enhance our understanding of resistance drift in the amorphous phase leading to improved multilevel memory operation in devices.

In this paper, we describe the development, validation and application of a "machine-learned" interatomic potential for a ternary system of practical relevance: namely, the $Ge_2Sb_2Te_5$ composition, which is used both in prototype and in commercial PCRAM devices. We employ the Gaussian Approximation Potential (GAP) framework[17–19] to fit the DFT potential-energy surfaces for a range of configurations, which has been successfully validated for a number of crystalline systems of increasing complexity[20–22] and is now becoming applicable to functional materials, such as carbonaceous electrodes for battery applications.[23,24] It was recently shown that the GAP approach can provide realistic and accurate large-scale models of amorphous solids,[25,26] but these previous studies have been limited to *elemental* systems—a limitation that is overcome in the present study of a ternary system for the first time. Additional information about the GAP potential developed here for $Ge_2Sb_2Te_5$, in particular the choice of combined descriptors for atomic environments, is provided in the Computational Details section.

The GAP potential for $Ge_2Sb_2Te_5$ is capable of simulating amorphous models that would be inaccessible to cubic-scaling DFT methods. Such large models (7,200 atoms in this paper) are needed to probe the medium- and extended-range order of the amorphous $Ge_2Sb_2Te_5$ structure, as they do not suffer from significant finite-size effects and they extend beyond the estimated structural correlation length of the material. This potential is able to provide statistics for meaningful investigations of chemical bonding through the fast generation of a database of smaller amorphous models and to simulate the crystallization of one of these models to the cubic phase while allowing the density to change at constant pressure and temperature.

Results and Discussion

Local Structure. Figure 1a shows images of structural models of amorphous ("*a*-") $Ge_2Sb_2Te_5$ generated using the GAP potential and produced by quenching from the melt. The models contain either 315 atoms (GAP_{315}) (comparable to the maximum model size achievable by DFT-based simulations) or 7,200 atoms (GAP_{7K}), the latter illustrating the capability of the O(*N*) GAP potential to create large models of this ternary phase in long MD simulations. To validate the quality of our GAP-generated structures, simulated neutron and X-ray diffraction patterns were calculated from the inverse Fourier transform of the radial-distribution functions (RDFs) and compared with previously reported experimental results.[28,29] For further validation, we use a previously reported DFT-MD trajectory for a 315-atom model of $Ge_2Sb_2Te_5$ (which had been obtained using a similar computational protocol, but was not part of the fitting); structural data for this are taken from ref. 27 and are hereafter labeled as "DFT_{315}". The GAP_{315} and DFT_{315} models agree well with each other, for both the neutron (Figure 1b) and X-ray structure factors (Figure S6 in the Supporting Information). Furthermore, the neutron-diffraction pattern of the large GAP_{7K} model shows a hint of the first sharp diffraction peak (FSDP) at $Q \sim 1.1$ Å^{-1}, and

hence a slightly better agreement with experiment compared to smaller models.[29] A FSDP has been observed in a similar region in both experimental and simulation studies of other amorphous chalcogenides, such as $GeSe_2$ (ref. 30) and $Ge_{15}Te_{85}$ (ref. 31). It has been argued that the FSDP is apparent only in the neutron-diffraction pattern due to Ge being a stronger scatterer of neutrons than Te (and also Sb), while the reverse is true for X-rays.[31] The presence of an appreciable FSDP only in the neutron-scattering pattern therefore provides strong evidence for the fact that "cation"-centered correlations are responsible for the FSDP, in accordance with the model proposed by Elliott.[32] The GAP_{7K} model shows only a modest FSDP, perhaps due to the fast quench rate (–15 K ps^{-1}). However, neither the DFT_{315} nor the GAP_{315} models generated with the same protocol show any sign of a FSDP, which indicates the importance of using a model size larger than the structural correlation length of *a*-$Ge_2Sb_2Te_5$, for the investigation of the medium-range order in this system. From an analysis of the long-range fluctuations in the RDF of the GAP_{7K} model, the correlation length of *a*-$Ge_2Sb_2Te_5$ is estimated to be about 20 Å (Figure S5).

As a further means of validation of the GAP potential, energy–volume equations of state were computed for the stable crystal structures of all relevant elemental (Ge, Sb, Te) and binary (GeTe, Sb_2Te_3) phases, as well as for layered, trigonal $Ge_2Sb_2Te_5$ in the Kooi stacking sequence.[33] Formation energies were calculated relative to the stable elemental phases at zero pressure. Correctly describing the energetics of competing phases is particularly important to avoid unphysical phase segregation in simulations of the ternary amorphous phase. Although the potential described herein has been developed for the technologically relevant $Ge_2Sb_2Te_5$ phase, our tests show that it does describe the constituent crystalline phases correctly as well (Figure 1c), indicating that it is also transferable to configurations sufficiently similar to those included in the training set.

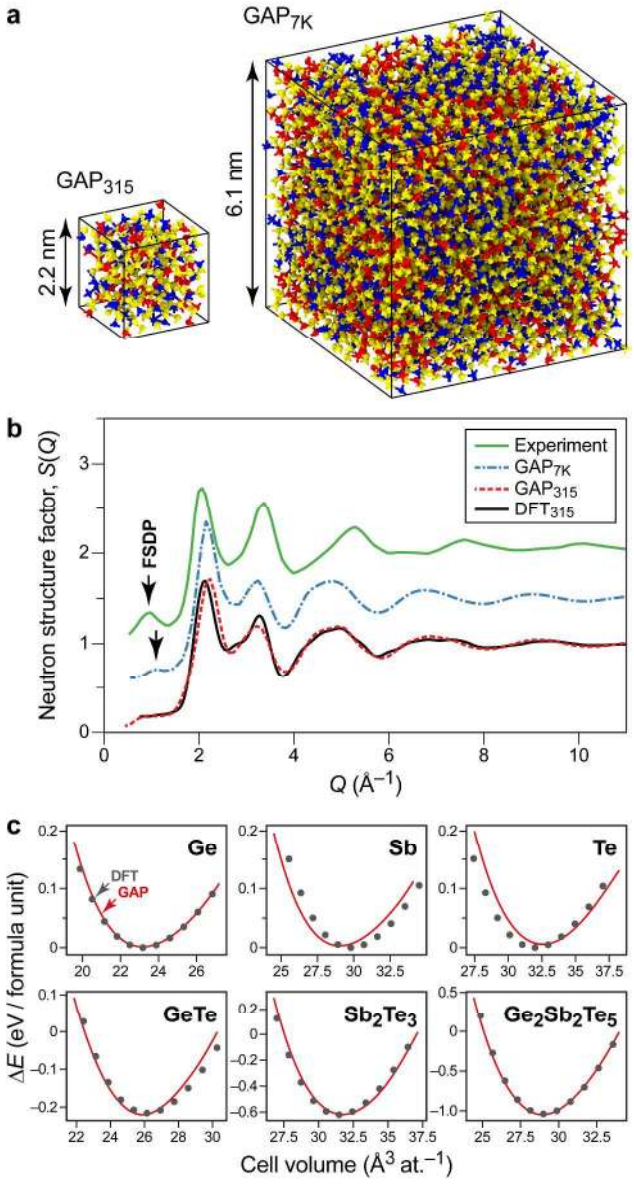


Figure 1. Large-scale structural modeling of *a*-$Ge_2Sb_2Te_5$ using the GAP machine-learning framework and validation of the method. (a) 315-atom and 7,200-atom GAP models of amorphous $Ge_2Sb_2Te_5$, visualized using OVITO.[34] Ge atoms are blue, Sb atoms red and Te atoms yellow. (b) Neutron structure factors of *a*-$Ge_2Sb_2Te_5$ from experiment, taken from ref. 29 (green), and calculated for various simulated models: DFT_{315} (black) from ref. 27, GAP_{315} (red) and GAP_{7K} (blue). The first sharp diffraction peak (FSDP) is highlighted by arrows; it is absent in the small GAP_{315} and DFT_{315} models. The experimental structure factor is shifted up by 1.0, and the GAP_{7K} structure factor is shifted up by 0.5, for clarity. (c) Equations of state for the crystalline elemental and binary constituents (in their respective ground-state structures) and for crystalline $Ge_2Se_2Te_5$ (Kooi stacking sequence). DFT reference data are shown as points, GAP results as red lines; formation energies are given relative to the respective DFT energies of the stable elemental phases at zero pressure.

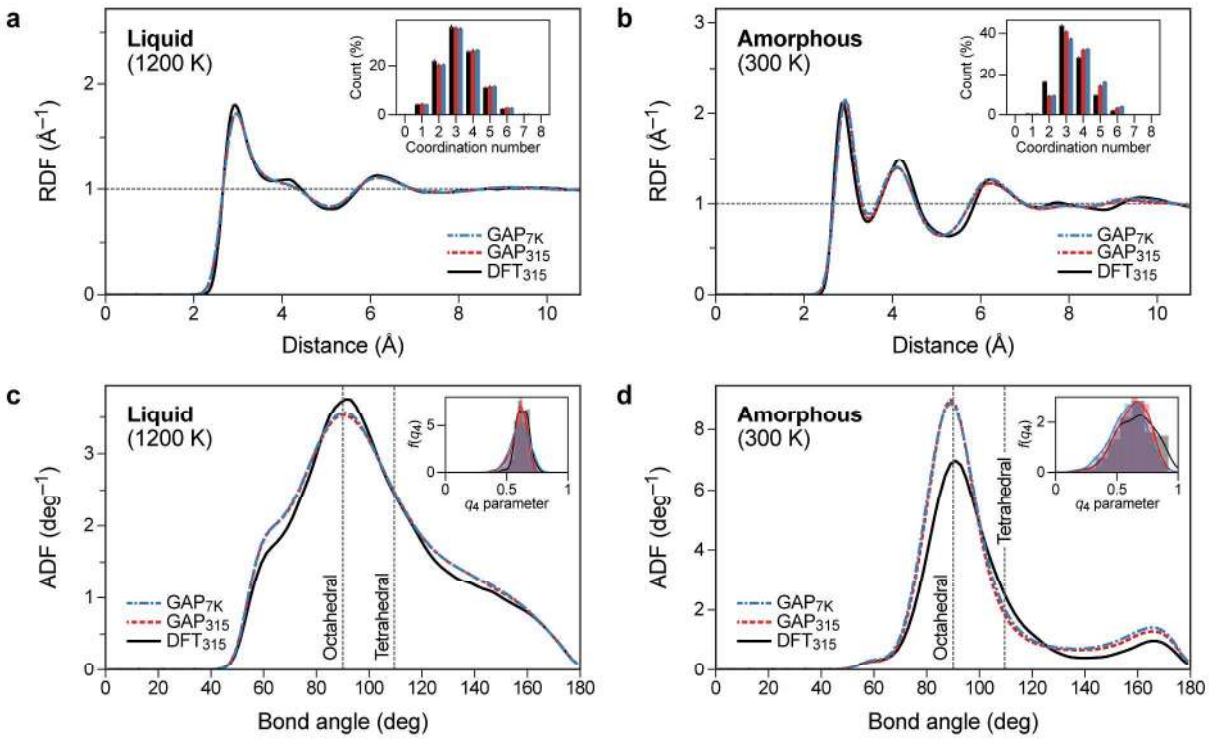


Figure 2. Total radial distribution functions (RDFs) and average coordination-number histograms (inset, calculated with a cut-off of 3.2 Å) for: (a) liquid; and (b) amorphous $Ge_2Sb_2Te_5$ models. Total angle-distribution functions (ADFs) and plots of the tetrahedral angular order-parameter, q_4, shown as insets for: (c) liquid; and (d) amorphous $Ge_2Sb_2Te_5$ models. The different simulated models are DFT (black) from ref. 27, GAP_{315} (red) and GAP_{7K} (blue). In the liquid phase, the GAP models are close to the DFT reference, while in the amorphous phase, the GAP models appear to be somewhat over-structured and have a slightly reduced fraction of tetrahedra.

The local structures of the liquid and amorphous phases were analyzed for the GAP_{315} and GAP_{7K} models. Total RDFs, coordination numbers calculated with a geometric bonding cut-off of 3.2 Å, and total angle-distribution functions (ADFs) are shown in Figure 2. Moreover, the distribution of the tetrahedral angular order parameter (q_4),[35] calculated in the manner of ref. 36, was used to estimate the fraction of tetrahedral-like environments in the different models (insets to Figures 2c and 2d).

An excellent agreement is observed for the liquid phase at 1200 K of $Ge_2Sb_2Te_5$ between the GAP models and DFT (Figures 2a and 2c). Broad and similar traces are observed for the total RDF and ADF of all simulated models with almost identical coordination numbers (inset of Figure 2a). The most probable coordination is 3-fold, and defective octahedral environments dominate, as evidenced by the broad peak at 90° in the ADF in Figure 2c. Very similar distributions are obtained also for the q_4 parameter of all the models (inset of Figure 2c). The proportion of tetrahedral-like environments in the liquid at 1200 K, as estimated from the integral of q_4 between 0.8 and 1.0 for the Ge- and Sb-centered environments,[37] is less than 1% for all models.

The short-range order of the amorphous phase for the GAP models at 300 K is analyzed in Figures 2b and 2d. The total RDFs exhibit a very good agreement with the DFT reference. A preference for 3-fold coordination is observed, as in the liquid, but with fewer 2-fold and almost no singly-coordinated atoms. The coordination numbers given in the inset of Figure 2b have a reduced proportion of 2- and 3-fold coordinated atoms and an increased proportion of 4- and 5-fold coordinated atoms for the GAP models compared to the DFT reference. There is again a peak in the ADF at 90°, now sharper than in the liquid. The GAP models also have a sharper ADF than the DFT reference, with a reduced density in the region corresponding to tetrahedral coordination. This suggests that the additional, relative to the DFT reference, 4- and 5-fold coordinated atoms present in the GAP models are part of defective octahedra. The q_4 trace given in the inset of Figure 2d is shifted to the right compared to the liquid for all models, revealing an increase of tetrahedral-like environments in the amorphous phase at 300 K compared to the liquid at 1200 K. The proportion of tetrahedral-like environments is smaller for the GAP_{315} (6–11 %) and GAP_{7K} (5 %) models compared to the DFT model (22 %). However, the question is far from fully resolved: other DFT models estimate the fraction of tetrahedral-like atoms in amorphous $Ge_2Sb_2Te_5$ or GeTe to be in the range 15–35 % (refs 12,37–39); furthermore, the inclusion of

dispersion interactions may increase the tetrahedral Ge fraction and percentage of "wrong" bonds,[40,41] as may the use of different pseudopotential constructions.[42] We stress that the GAP models generated in this work aim to reproduce the reference DFT structures generated at the same computational level,[27] and appear to be only slightly over-structured compared to the latter.

The reduced percentage of tetrahedral-like local environments can be possibly linked to a lack of so-called "wrong" bonds, more precisely bonds that are not present in the ideal cubic (rocksalt) crystal structure. These are: Ge−Ge, Ge−Sb, Sb−Sb and Te−Te bonds.[43] The statistical scarcity of these bonds may be inferred from the reduced intensity in the first peak of the partial RDFs for these atom pairs in the GAP models (Figures S3–S4). In particular, the homopolar Ge−Ge bond is known to stabilize tetrahedral environments but not octahedral ones.[44] Bonding defects, such as the "wrong" bonds and tetrahedral environments, were under-sampled in the training set, which consists of predominantly high-temperature liquid, and annealed amorphous, structures. Therefore, the forces that drive configurations away from the thermally unfavorable arrangements containing tetrahedra and "wrong" bonds could be overestimated by this GAP model during the MD trajectory of the non-equilibrium quench. Additional training data, such as an alchemical sampling of antisite configurations,[45] samples from lower temperatures or higher pressures, as well as samples with different compositions in the ternary phase diagram, could further improve the potential. This will be the topic of future work.

Medium-range order. The medium-range order of the GAP models was characterized by calculating the statistics of shortest-path rings, consisting of up to 27 atoms, using the Franzblau algorithm.[46] A uniform geometric cut-off of 3.2 Å was used to define bonded pairs, and the resulting network of atoms has a graph diameter (longest node-to-node path in the atomic-network graph) of ∼18 in the case of 315-atom models and ∼36 for the 7,200-atom models. The statistics

of shortest-path rings in the GAP models agree with the DFT models, as shown in Figure 3a (liquid phase) and Figure 3c (amorphous phase).

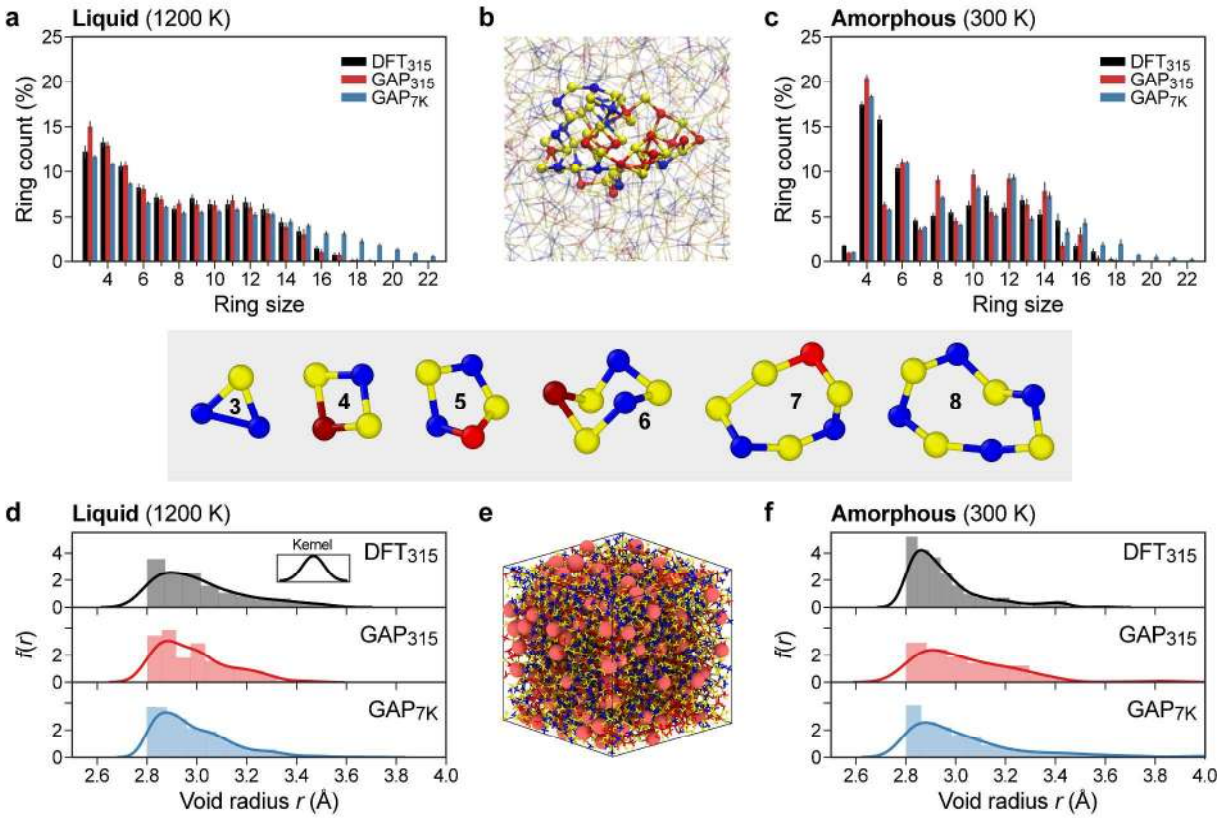


Figure 3. (a–c) Medium-range order in $Ge_2Sb_2Te_5$ as analyzed through ring statistics. The distribution of shortest-path rings in liquid (panel a) and amorphous (panel c) $Ge_2Sb_2Te_5$ models is shown for DFT (black) from ref. 27, GAP_{315} (red) and GAP_{7K} (blue). Panel (b) shows a cluster of interconnected shortest-path rings from the GAP_{7K} model of *a*-$Ge_2Sb_2Te_5$, revealing the presence of significant alternation of A (Ge (blue), Sb (red)) and B (Te (yellow)) atoms; the shaded area below shows examples of shortest-path rings from the amorphous structure. (d–f) Voids in $Ge_2Sb_2Te_5$. Distributions of void radii are shown for different liquid (panel d) and amorphous (panel f) $Ge_2Sb_2Te_5$ models; solid lines are kernel-density estimates[47,48] of the void-radius probability distributions using a Gaussian kernel, shown as an inset. Panel (e) shows the amorphous GAP_{7K} model, where voids are highlighted as pink spheres of largest radius.

In the liquid at 1200 K, very similar distributions of shortest-path rings can be observed for the different models. Small rings, 3−6 fold, are entropically favored and have the largest relative

probabilities in the equilibrium liquid phase. Rings of intermediate size, 7−12 fold, have an almost uniform distribution. Beyond that, for the GAP_{7K} model, the distribution decays relatively slowly with ring size, revealing the existence of rings consisting of up to 27 atoms that cannot be observed in the smaller models but which are below the calculated graph diameter (Figure S7).

In the amorphous phase, the GAP models capture the previously reported pattern of an increased probability of even-membered rings over odd-membered rings.[38,39] A cluster of interconnected rings from the GAP_{7K} model, which contains the atom present in the most rings, is shown in Figure 3b. There is a visible quasi-binary alternation of A (Ge, Sb) and B (Te) atoms in the amorphous phase, resulting from a significant statistical preference for so-called "correct"-bonds, namely A−B heteropolar bonds which form the ideal crystalline cubic phase, viz. Ge−Te and Sb−Te bonds.[28,38,39] Even-membered rings can have perfect A−B alternation while odd rings must have at least one "wrong"-bond in their structure (either A−A or B−B). The enthalpic penalty for having a "wrong"-bond is most important for the 3-fold rings and results in a drastically reduced population of this ring size in the amorphous phase at 300 K compared to the liquid phase at 1200 K. The four-fold ring is the most probable motif in the amorphous phase and is a building block for crystallization to the metastable cubic phase, where four-fold square rings are the characteristic topological unit.[4,38] A reduced population of five-fold rings is observed in the GAP models compared to the reference DFT model which does not have a perfect odd/even ring alternation. However, other experimental data and some of the best available simulation work, including models that are larger than our current DFT reference, suggest that the probability of five-fold rings in the amorphous phase could be lower than the corresponding probability of both four- and six-fold rings and that shortest-path ring odd-even alternation is likely to be pronounced in *a*-$Ge_2Sb_2Te_5$.[28,38]

Vacancies in the crystal, and correspondingly voids in the amorphous phase, are a common occurrence in $Ge_2Sb_2Te_5$ and play an important role in its structure, dynamics and electronic properties.[49–52] The interplay between voids, rings and lone pairs influences the dynamics of crystallization.[51] In this work, voids are defined on a fine mesh with a maximal spacing of 0.3 Å using the Voronoi construction with a cut-off of 2.8 Å, such that there are no interstitial voids in the ideal cubic crystal.[51] The distributions of the radii of largest spheres inside these voids for liquid and amorphous $Ge_2Sb_2Te_5$ are shown in Figures 3d and 3f, and a snapshot of such spheres within the GAP_{7K} model is presented in Figure 3e.

Chemical Bonding. A crucial question is whether the GAP developed in this work is able to reproduce the complex distribution of local environments present in *a*-$Ge_2Sb_2Te_5$.[53] Previous work has shown that the "fluxional" (rapidly interconverting at elevated temperature) local environments in $Ge_2Sb_2Te_5$ are intimately connected with the lone pairs present on many of the atoms in the system and the establishment of a bonding-energy hierarchy.[27,54,55] To distinguish local environments, the typical geometric approach based on a bond-distance cut-off is insufficient, as there is no satisfactory choice that does not leave too many weak bonds outside of the coordination sphere or otherwise introduces non-bonded atoms into it. An optimal geometric cut-off, that minimizes the number of errors, can be inferred from charge-density information, as obtained from an electronic-structure calculation in several different ways.[31,44] A value for this cut-off between 3.0 and 3.2 Å is typically chosen for this material. Nevertheless, whenever resources allow, identifying bonds directly from the charge-density information is preferable.[27]

The electron-localization function (ELF) was introduced to DFT by Savin et al.[56] Through topological analysis of the attractors in the ELF scalar field, evidence can be gathered about the existence of covalent bonds and lone pairs.[57] The local environments revealed by this analysis are in agreement with valence-shell electron-pair repulsion (VSEPR) theory. DFT investigations of

the ELF in amorphous or liquid models of materials such as $Ge_2Sb_2Te_5$ require statistical averaging over many configurations to adequately describe the system, which is not easily possible to achieve with DFT. By using the GAP first to generate and sample numerous independent liquid and amorphous structures at a much reduced computational cost, and then feeding these into DFT electronic-structure calculations, these investigations, indeed any DFT investigation that requires accurate structural input of this sort, can be enhanced and made more reproducible.[24] Here, the GAP was used to generate five uncorrelated *a*-$Ge_2Sb_2Te_5$ 315-atom structural models, which is easily possible as simulation with the GAP is two orders of magnitude faster than for DFT for an equivalent MD protocol at this system size. Amorphous GAP models were cooled down to 0 K and had their geometry optimized with the GAP before they were fed into DFT calculations for subsequent ELF analysis and comparison with previous results from DFT-sampled structures.[27] Figure 4a reports the coordination numbers obtained by imposing a cut-off for the minimum ELF value in a basin between two atoms ($\min_{\mathbf{r}} \eta(\mathbf{r}) < \eta_{cut}$). Coordination numbers for the GAP models, as obtained from the ELF analysis, correspond well with those of the reference DFT model: three- or four-fold coordination is the favorable arrangement for Ge local environments, while three-fold coordinated environments are the most probable for Sb and Te atoms. Compared to the typical bond-length cut-off analysis, the ELF discrimination removes unphysical local environments. Figure 4b shows the observed majority of bonding geometries for Ge in the GAP models, which agree very well with the reported structural units from DFT calculations.[28,29,37,38] In particular, 4-fold coordinated environments, such as the tetrahedral and see-saw structures, can be distinguished in the amorphous structure of $Ge_2Sb_2Te_5$. The latter defective-octahedral geometry involves two long axial and two short equatorial bonds, whose characteristics are successfully reproduced by the GAP models; for instance, the bond lengths for axial and equatorial bonds are

2.94−3.04 Å and 2.80−2.82 Å, respectively. Compared to these, the tetrahedral bond length is much shorter, at 2.62−2.67 Å, again in agreement with the DFT study.[27]

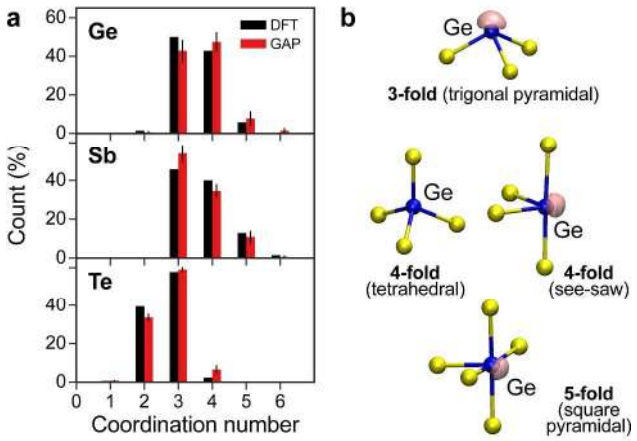


Figure 4. (a) Coordination-number histograms obtained from an ELF analysis for Ge, Sb and Te atoms. Shown in black are results from DFT simulation, from ref. 27, and in red are the averaged results from five different 315-atom fully quenched amorphous models generated with GAP; error bars reveal a small model-to-model variation. Representative local environments of Ge (in blue) surrounded by Te atoms (in yellow) are given in (b); these are present in both the DFT model as well as in the five different GAP models of amorphous $Ge_2Sb_2Te_5$. To highlight the presence of lone pairs, an ELF isosurface (mauve) with values in the range 0.85−0.9 is shown for appropriate configurations.

Crystallization. Equally important is the ability of the potential to capture the fast transition of *a*-$Ge_2Sb_2Te_5$ to the metastable cubic phase during annealing. In order to investigate this, one of the 315-atom GAP models of *a*-$Ge_2Sb_2Te_5$, generated for the ELF analysis, was slowly heated (15 K ps^{-1}) and then annealed at 600 K, at constant zero pressure, for up to 800 ps to study crystallization to the metastable cubic phase. The transformation was monitored by a Fourier analysis of the density fluctuations.[4] Furthermore, to quantify the increase in chemical and topological order, the number of 4-fold shortest-path rings with quasi-binary AB alternation was calculated during the trajectory and normalized against the expected number of rings in an ideal crystal. Results are summarized in Figure 5 and reveal the capability of the GAP to capture the crystalli-

zation of *a*-$Ge_2Sb_2Te_5$. By allowing the volume to change during the simulation, finite-size effects are partially mitigated. Details of the resulting density changes during the crystallization anneal are provided in the SI (Figure S9). The crystallization of additional models including larger ones needs to be carried out in order to compare quantitatively with experiment and this will be the focus of future work.

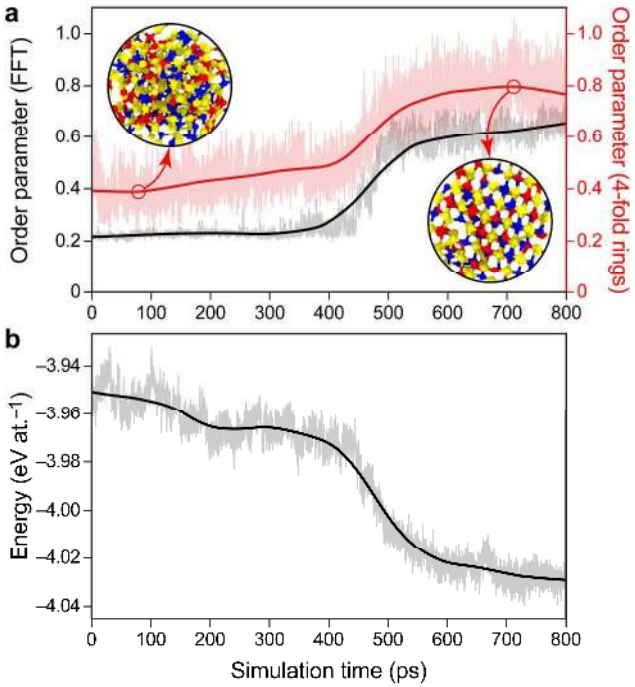


Figure 5. Evolution during a 600 K constant-pressure anneal of one of the GAP 315-atom *a*-$Ge_2Sb_2Te_5$ models. (a) The second largest component of the 3D Fourier transform of the atomic density is shown in black (*y*-axis on the left), while the ratio of the number of shortest-path four-fold rings with AB alternation, N_{AB}, over the number in the ideal cubic crystal is shown in red (*y*-axis on the right). The structures with the minimal and maximal four-fold ring order are marked with red circles on the corresponding order-parameter line and images of these structures are displayed as insets. (b) Potential energy predicted by the GAP. The structure becomes more chemically and topologically ordered during the anneal, while the energy decreases. Solid lines are locally weighted smoothing traces,[58] while the raw data are shown with a slight transparency.

Conclusions

The GAP machine-learning framework has been used to fit, for the first time, the potential-energy surface of the ternary $Ge_2Sb_2Te_5$ phase-change memory material. This is a significant milestone toward the routine and realistic modeling of chemically complex functional materials. Structural models of liquid and amorphous $Ge_2Sb_2Te_5$ obtained from GAP-MD agree well with DFT-MD simulations and with experimental diffraction data. The GAP model with 7,200 atoms begins to capture the first-sharp diffraction peak in the neutron-diffraction pattern, which is not visible in the simulated patterns for smaller models. This reveals the necessity for generating large-scale structures to investigate the medium- and extended-range order of amorphous $Ge_2Sb_2Te_5$—which are currently out of reach for DFT, but readily within reach for GAP. We also generated several uncorrelated, 315-atom models of *a*-$Ge_2Sb_2Te_5$ and analyzed their bonding nature. By such studies, GAP can be used to enhance statistical reproducibility by rapidly generating databases of small structural models that can then be used for advanced electronic-structure calculations. Finally, GAP-MD captures the crystallization of the amorphous phase, including the change in density. Future directions of this work include the study of dynamical and thermal properties of the supercooled liquid and amorphous phases of $Ge_2Sb_2Te_5$, surfaces and superlattices (through suitable extensions of the "training" database), and the application of the concepts developed here to GST phase-change materials with different chemical compositions.

Computational Details

Sampling and Evaluation. The GAP interatomic potential for $Ge_2Sb_2Te_5$ is fitted to a reference database of DFT data, for which appropriate sampling of structures and local atomic environments is crucial. The most important part of the reference structures was collected from DFT-MD trajectories for $Ge_2Sb_2Te_5$ in the liquid, amorphous and partially crystallized states (Supporting Information); for this task, we used the projector augmented-wave method[59] as implemented

in VASP[60,61] with settings as in ref. 27. Additionally, reference structures were added from several iterations of Monte Carlo sampling for the primitive unit cells of six stable compositions in the ternary phase diagram (cf. Figure 1c), so as to correct the chemical-potential dependence and to prevent non-physical phase separation (or demixing) at low temperatures.

Single-point energy, force, and virial stress data were then calculated for representative structures in the database, again using VASP,[60,61] but employing much more stringent and well-converged settings (e.g. a plane-wave cutoff of 600 eV and a *k*-point density of 0.016 $Å^{-1}$) than what would, or could, be used in a typical DFT-MD simulation (often, these use a < 300 eV cutoff, and only a single *k*-point). We found such high accuracy to be important to minimize noise in the training input and to ensure accurate forces and stresses.

Training. The interatomic potential was generated using the GAP ML framework[17,18] as implemented in QUIP (http://www.libatoms.org); the general approach has been reviewed in ref. 19. To prevent non-physical clustering of atoms at a high temperature, we use, akin to ref. 62, a linear combination of two descriptors for the atomic structure: a non-parametric two-body term and the Smooth Overlap of Atomic Positions (SOAP).[18] The Supporting Information provides further details on the composition of the reference database and the specific program input used for the fit.

GAP-Driven MD. Once the GAP was finalized, it was used for model-generation simulations as follows. GAP-MD simulations were carried out with the LAMMPS code,[63] which is directly interfaced to QUIP and freely available at http://lammps.sandia.gov. The melt-and-quench protocol, used both for DFT-MD and GAP-MD, is illustrated in Figure S1; canonical ensemble (NVT) MD simulations were performed in a cubic box with periodic boundary conditions at the experimental amorphous density of 5.88 g cm^{-3} (refs. 64, 65). This protocol is consistent with the one used for the DFT reference data from our previous work.[27] The timestep was 1 fs. Furthermore, GAP-MD annealing simulations of crystallization (cf. Figure 5) were carried out in an isobaric-isothermal (NPT) ensemble to allow the volume to change naturally during crystallization. The density increased from 5.88 g cm^{-3} at the start of the anneal to around 6.15 g cm^{-3} for the configuration with the maximum crystallinity (FFT) order parameter. This is in reasonable agreement with the experimental density of the metastable cubic crystal phase of about 6.3 g cm^{-3}.[64,65]

Acknowledgements

F.C.M. thanks A. P. Bartók for helpful discussions. F.C.M also acknowledges financial support from the UK Engineering and Physical Sciences Research Council (EPSRC) Centre for Doctoral Training in Computational Methods for Materials Science under grant EP/L015552/1. V.L.D. acknowledges a Feodor Lynen Research 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. This work used the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk) via the UKCP Consortium (EP/K013564/1) and Materials Chemistry Consortium (EP/L000202), both of which are funded by the EPSRC.

Associated Content

Supporting Information. Detailed supplementary information regarding: GAP fitting, database composition, and MD simulations; detailed supplementary results regarding atomic structure, including the long-range decay of correlations in the RDF.

Data access statement: Data supporting this publication will be made available through an online repository after acceptance.

References

(1) Raoux, S.; Xiong, F.; Wuttig, M.; Pop, E. Phase Change Materials and Phase Change Memory. *MRS Bull.* **2014**, *39*, 703–710.

(2) Wuttig, M.; Yamada, N. Phase-Change Materials for Rewriteable Data Storage. *Nat. Mater.* **2007**, *6*, 824–832.

(3) Deringer, V. L.; Dronskowski, R.; Wuttig, M. Microscopic Complexity in Phase-Change Materials and its Role for Applications. *Adv. Funct. Mater.* **2015**, *25*, 6343–6359.

(4) Hegedüs, J.; Elliott, S. R. Microscopic Origin of the Fast Crystallization Ability of Ge–Sb–Te Phase-Change Memory Materials. *Nat. Mater.* **2008**, *7*, 399–405.

(5) Hegedus, J.; Elliott, S. R. Computer□Simulation Design of New Phase□change Memory Ma-

terials. *phys. stat. sol.* **2010**, *207*, 510–515.

(6) Lee, T. H.; Elliott, S. R. *Ab Initio* Computer Simulation of the Early Stages of Crystallization: Application to $Ge_2Sb_2Te_5$ Phase-Change Materials. *Phys. Rev. Lett.* **2011**, *107*, 145702.

(7) Zhang, W.; Deringer, V. L.; Dronskowski, R.; Mazzarello, R.; Ma, E.; Wuttig, M. Density-Functional Theory Guided Advances in Phase-Change Materials and Memories. *MRS Bull.* **2015**, *40*, 856–869.

(8) Kalikka, J.; Akola, J.; Jones, R. O. Crystallization Processes in the Phase Change Material $Ge_2Sb_2Te_5$: Unbiased Density Functional/Molecular Dynamics Simulations. *Phys. Rev. B* **2016**, *94*, 134105.

(9) Zipoli, F.; Curioni, A. Reactive Potential for the Study of Phase-Change Materials: GeTe. *New J. Phys.* **2013**, *15*, 123006.

(10) Zipoli, F.; Krebs, D.; Curioni, A. Structural Origin of Resistance Drift in Amorphous GeTe. *Phys. Rev. B* **2016**, *93*, 115201.

(11) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(12) Sosso, G. C.; Miceli, G.; Caravati, S.; Behler, J.; Bernasconi, M. Neural Network Interatomic Potential for the Phase Change Material GeTe. *Phys. Rev. B* **2012**, *85*, 174103.

(13) Sosso, G. C.; Miceli, G.; Caravati, S.; Giberti, F.; Behler, J.; Bernasconi, M. Fast Crystallization of the Phase Change Compound GeTe by Large-Scale Molecular Dynamics Simulations. *J. Phys. Chem. Lett.* **2013**, *4*, 4241–4246.

(14) Sosso, G. C.; Colombo, J.; Behler, J.; Del Gado, E.; Bernasconi, M. Dynamical Heterogeneity in the Supercooled Liquid State of the Phase Change Material GeTe. *J. Phys. Chem. B* **2014**, *118*, 13621–13628.

(15) Campi, D.; Donadio, D.; Sosso, G. C.; Behler, J.; Bernasconi, M. Electron-Phonon Interaction and Thermal Boundary Resistance at the Crystal-Amorphous Interface of the Phase Change Compound GeTe. *J. Appl. Phys.* **2015**, *117*, 015304.

(16) Gabardi, S.; Baldi, E.; Bosoni, E.; Campi, D.; Caravati, S.; Sosso, G. C.; Behler, J.; Bernasconi, M. Atomistic Simulations of the Crystallization and Aging of GeTe Nanowires. *J. Phys. Chem. C* **2017**, *121*, 23827–23838.

(17) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.

(18) Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B* **2013**, *87*, 184115.

(19) Bartók, A. P.; Csányi, G. Gaussian Approximation Potentials: A Brief Tutorial Introduction. *Int. J. Quantum Chem.* **2015**, *115*, 1051–1057.

(20) Bartók, A. P.; De, S.; Poelking, C.; Bernstein, N.; Kermode, J. R.; Csányi, G.; Ceriotti, M. Machine Learning Unifies the Modeling of Materials and Molecules. *Sci. Adv.* **2017**, *3*, e1701816.

(21) Dragoni, D.; Daff, T. D.; Csányi, G.; Marzari, N. Achieving DFT Accuracy with a Machine-Learning Interatomic Potential: Thermomechanics and Defects in bcc Ferromagnetic Iron. *Phys. Rev. Mater.* **2018**, *2*, 013808.

(22) Deringer, V. L.; Pickard, C. J.; Csányi, G. Data-Driven Learning of Total and Local Energies in Elemental Boron. *Phys. Rev. Lett.* **2018**, *120*, 156001.

(23) Fujikake, S.; Deringer, V. L.; Lee, T. H.; Krynski, M.; Elliott, S. R.; Csányi, G. Gaussian Approximation Potential Modeling of Lithium Intercalation in Carbon Nanostructures. *J. Chem. Phys.* **2018**, *148*, 241714.

(24) Deringer, V. L.; Merlet, C.; Hu, Y.; Lee, T. H.; Kattirtzi, J. A.; Pecher, O.; Csányi, G.; Elliott, S. R.; Grey, C. P. Towards an Atomistic Understanding of Disordered Carbon Electrode Materials. *Chem. Commun.* **2018**, *54*,5988-5991.

(25) Caro, M. A.; Deringer, V. L.; Koskinen, J.; Laurila, T.; Csányi, G. Growth Mechanism and Origin of High sp^3 Content in Tetrahedral Amorphous Carbon. *Phys. Rev. Lett.* **2018**, *120*, 166101.

(26) Deringer, V. L.; Bernstein, N.; Bartók, A. P.; Kerber, R. N.; Cliffe, M. J.; Marbella, L. E.; Grey, C. P.; Elliott, S. R.; Csányi, G. Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics. *J. Phys. Chem. Lett.* **2018**, *9*, 2879–2885.

(27) Lee, T. H.; Elliott, S. R. The Relation between Chemical Bonding and Ultrafast Crystal Growth. *Adv. Mater.* **2017**, *29*, 1700814.

(28) Kohara, S.; Kato, K.; Kimura, S.; Tanaka, H.; Usuki, T.; Suzuya, K.; Tanaka, H.; Moritomo,

Y.; Matsunaga, T.; Yamada, N.; et al. Structural Basis for the Fast Phase Change of $Ge_2Sb_2Te_5$: Ring Statistics Analogy between the Crystal and Amorphous States. *Appl. Phys. Lett.* **2006**, *89*, 201910.

(29) Jóvári, P.; Kaban, I.; Steiner, J.; Beuneu, B.; Schöps, A.; Webb, M. A. Local Order in Amorphous $Ge_2Sb_2Te_5$ and $GeSb_2Te_4$. *Phys. Rev. B* **2008**, *77*, 035202.

(30) Vashishta, P.; Kalia, R. K.; Antonio, G. A.; Ebbsjö, I. Atomic Correlations and Intermediate-Range Order in Molten and Amorphous $GeSe_2$. *Phys. Rev. Lett.* **1989**, *62*, 1651–1654.

(31) Kalikka, J.; Akola, J.; Jones, R. O.; Kohara, S.; Usuki, T. Amorphous $Ge_{15}Te_{85}$: Density Functional, High-Energy x-Ray and Neutron Diffraction Study. *J. Phys. Condens. Matter* **2012**, *24*, 015802.

(32) Elliott, S. R. Origin of the First Sharp Diffraction Peak in the Structure Factor of Covalent Glasses. *Phys. Rev. Lett.* **1991**, *67*, 711–714.

(33) Kooi, B. J.; De Hosson, J. T. M. Electron Diffraction and High-Resolution Transmission Electron Microscopy of the High Temperature Crystal Structures of $Ge_xSb_2Te_{3+x}$ (x = 1,2,3) Phase Change Material. *J. Appl. Phys.* **2002**, *92*, 3584–3590.

(34) Stukowski, A. Visualization and Analysis of Atomistic Simulation Data with OVITO–the Open Visualization Tool. *Model. Simul. Mater. Sci. Eng.* **2010**, *18*, 015012.

(35) Chau, P.-L.; Hardwick, A. J. A New Order Parameter for Tetrahedral Configurations. *Mol. Phys.* **1998**, *93*, 511–518.

(36) Errington, J. R.; Debenedetti, P. G. Relationship between Structural Order and the Anomalies of Liquid Water. *Nature* **2001**, *409*, 318–321.

(37) Caravati, S.; Bernasconi, M.; Kühne, T. D.; Krack, M.; Parrinello, M. Coexistence of Tetrahedral- and Octahedral-like Sites in Amorphous Phase Change Materials. *Appl. Phys. Lett.* **2007**, *91*, 171906.

(38) Akola, J.; Jones, R. O. Structural Phase Transitions on the Nanoscale: The Crucial Pattern in the Phase-Change Materials $Ge_2Sb_2Te_5$ and GeTe. *Phys. Rev. B* **2007**, *76*, 235201.

(39) Akola, J.; Jones, R. O. Density Functional Study of Amorphous, Liquid and Crystalline $Ge_2Sb_2Te_5$: Homopolar Bonds and/or AB Alternation? *J. Phys. Condens. Matter* **2008**, *20*, 465103.

(40) Micoulaut, M. Communication: Van Der Waals Corrections for an Improved Structural Description of Telluride Based Materials. *J. Chem. Phys.* **2013**, *138*, 061103.

(41) Micoulaut, M.; Piarristeguy, A.; Flores-Ruiz, H.; Pradel, A. Towards Accurate Models for Amorphous GeTe: Crucial Effect of Dispersive van Der Waals Corrections on the Structural Properties Involved in the Phase-Change Mechanism. *Phys. Rev. B* **2017**, *96*, 184204.

(42) Bouzid, A.; Ori, G.; Boero, M.; Lampin, E.; Massobrio, C. Atomic-Scale Structure of the Glassy $Ge_2Sb_2Te_5$ Phase Change Material: A Quantitative Assessment via First-Principles Molecular Dynamics. *Phys. Rev. B* **2017**, *96*, 224204.

(43) Jóvári, P.; Kaban, I.; Steiner, J.; Beuneu, B.; Schöps, A.; Webb, A. "Wrong Bonds" in Sputtered Amorphous $Ge_2Sb_2Te_5$. *J. Phys. Condens. Matter* **2007**, *19*, 335212.

(44) Deringer, V. L.; Zhang, W.; Lumeij, M.; Maintz, S.; Wuttig, M.; Mazzarello, R.; Dronskowski, R. Bonding Nature of Local Structural Motifs in Amorphous GeTe. *Angew. Chem. Int. Ed.* **2014**, *53*, 10817–10820.

(45) Raty, J.-Y.; Zhang, W.; Luckas, J.; Chen, C.; Mazzarello, R.; Bichara, C.; Wuttig, M. Aging Mechanisms in Amorphous Phase-Change Materials. *Nat. Commun.* **2015**, *6*, 7467.

(46) Franzblau, D. S. Computation of Ring Statistics for Network Models of Solids. *Phys. Rev. B* **1991**, *44*, 4925–4930.

(47) Rosenblatt, M. Remarks on Some Nonparametric Estimates of a Density Function. *Ann. Math. Stat.* **1956**, *27*, 832–837.

(48) Parzen, E. On Estimation of a Probability Density Function and Mode. *Ann. Math. Stat.* **1962**, *33*, 1065–1076.

(49) Kolobov, A. V; Haines, J.; Pradel, A.; Ribes, M.; Fons, P.; Tominaga, J.; Steimer, C.; Aquilanti, G.; Pascarelli, S. Pressure-Induced Amorphization of Quasibinary GeTe–Sb_2Te_3: The Role of Vacancies. *Appl. Phys. Lett.* **2007**, *91*, 021911.

(50) Wuttig, M.; Lüsebrink, D.; Wamwangi, D.; Wełnic, W.; Gilleßen, M.; Dronskowski, R. The Role of Vacancies and Local Distortions in the Design of New Phase-Change Materials. *Nat. Mater.* **2007**, *6*, 122–128.

(51) Lee, T. H.; Elliott, S. R. Structural Role of Vacancies in the Phase Transition of $Ge_2Sb_2Te_5$ Memory Materials. *Phys. Rev. B* **2011**, *84*, 094124.

(52) Zhang, W.; Thiess, A.; Zalden, P.; Zeller, R.; Dederichs, P. H.; Raty, J.-Y.; Wuttig, M.; Blügel, S.; Mazzarello, R. Role of Vacancies in Metal–insulator Transitions of Crystalline Phase-Change Materials. *Nat. Mater.* **2012**, *11*, 952–956.

(53) Krbal, M.; Kolobov, A. V.; Fons, P.; Tominaga, J.; Elliott, S. R.; Hegedus, J.; Uruga, T. Intrinsic Complexity of the Melt-Quenched Amorphous $Ge_2Sb_2Te_5$ Memory Alloy. *Phys. Rev. B* **2011**, *83*, 054203.

(54) Kolobov, A. V.; Fons, P.; Tominaga, J.; Ovshinsky, S. R. Vacancy-Mediated Three-Center Four-Electron Bonds in GeTe-Sb_2Te_3 Phase-Change Memory Alloys. *Phys. Rev. B* **2013**, *87*, 165206.

(55) Kolobov, A. V.; Fons, P.; Tominaga, J. Athermal Amorphization of Crystallized Chalcogenide Glasses and Phase-Change Alloys. *phys. stat. sol.* **2014**, *251*, 1297–1308.

(56) Savin, A.; Becke, A. D.; Flad, J.; Nesper, R.; Preuss, H.; von Schnering, H. G. A New Look at Electron Localization. *Angew. Chem. Int. Ed.* **1991**, *30*, 409–412.

(57) Silvi, B.; Savin, A. Classification of Chemical Bonds Based on Topological Analysis of Electron Localization Functions. *Nature* **1994**, *371*, 683–686.

(58) Cleveland, W. S. Robust Locally Weighted Regression and Smoothing Scatterplots. *J. Am. Stat. Assoc.* **1979**, *74*, 829–836.

(59) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953–17979.

(60) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169–11186.

(61) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59*, 1758–1775.

(62) Deringer, V. L.; Csányi, G. Machine Learning Based Interatomic Potential for Amorphous Carbon. *Phys. Rev. B* **2017**, *95*, 094203.

(63) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.

(64) Nonaka, T.; Ohbayashi, G.; Toriumi, Y.; Mori, Y.; Hashimoto, H. Crystal Structure of GeTe and $Ge_2Sb_2Te_5$ Meta-Stable Phase. *Thin Solid Films* **2000**, *370*, 258–261.

(65) Njoroge, W. K.; Wöltgens, H.-W.; Wuttig, M. Density Changes upon Crystallization of $Ge_2Sb_{2.04}Te_{4.74}$ Films. *J. Vac. Sci. Technol. A* **2002**, *20*, 230–233.