

Ab initio phase diagram and nucleation of gallium

Haiyang Niu ^{1,2,3}✉, Luigi Bonati ^{3,4}, Pablo M. Piaggi^{2,3} & Michele Parrinello ^{2,3,5}✉

Elemental gallium possesses several intriguing properties, such as a low melting point, a density anomaly and an electronic structure in which covalent and metallic features coexist. In order to simulate this complex system, we construct an ab initio quality interaction potential by training a neural network on a set of density functional theory calculations performed on configurations generated in multithermal-multibaric simulations. Here we show that the relative equilibrium between liquid gallium, α -Ga, β -Ga, and Ga-II is well described. The resulting phase diagram is in agreement with the experimental findings. The local structure of liquid gallium and its nucleation into α -Ga and β -Ga are studied. We find that the formation of metastable β -Ga is kinetically favored over the thermodynamically stable α -Ga. Finally, we provide insight into the experimental observations of extreme undercooling of liquid Ga.

¹State Key Laboratory of Solidification Processing, International Center for Materials Discovery, School of Materials Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, P. R. China. ²Department of Chemistry and Applied Biosciences, ETH Zurich c/o USI Campus, Via Giuseppe Buffi 13, 6900 Lugano, Switzerland. ³Facoltà di Informatica, Istituto di Scienze Computazionali, and National Center for Computational Design and Discovery of Novel Materials MARVEL, Università della Svizzera italiana (USI), Via Giuseppe Buffi 13, 6900 Lugano, Switzerland. ⁴Department of Physics, ETH Zurich, c/o Università della Svizzera italiana, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland. ⁵Istituto Italiano di Tecnologia, Via Morego 30, 16163 Genova, Italy. ✉email: haiyang.niu@nwpu.edu.cn; parrinello@phys.chem.ethz.ch

Elemental gallium is a unique metal with a number of fascinating and unusual properties^{1–6}. In its different phases it finds a variety of important technological applications^{1,7}. Unlike most metals, including those in the same column of the periodic table, it crystallizes in rather complex structures^{5,8} as illustrated in Fig. 1. Its stable solid phase at ambient conditions, known as α -Ga, is orthorhombic with four atoms in the primitive unit cell⁸. Each atom is coordinated to seven neighbors, resulting into a highly anisotropic atomic environment³ (see Fig. 1a). The bonding between the two nearest neighbor atoms is described as covalent and α -Ga is thought to have a mixture of covalent and metallic bonds.

Upon melting, like water, it exhibits a density anomaly with the liquid expanding as much as 3.1%. Despite decades of experimental investigations, the structure of liquid gallium is not yet fully understood^{2,6}. Furthermore, micrometer-sized or submicrometer-size liquid gallium could be undercooled down to 150 K without solidification^{4,9–11}. In such scenario, the crystallization does not produce the stable α phase but mostly the β -Ga structure. Experiments also showed that Ga nanoparticles in the range of 3–15 nm can be undercooled even down to 90 K with no sign of a freezing transition¹². The structure of β -Ga is monoclinic and contains square-like arrangements of atoms forming layers parallel to the (021) plane as shown in Fig. 1b. At variance with α -Ga, each atom in β -Ga has eight nearest neighbors at the distance of 2.78 Å. In addition the Ga phase diagram exhibits a variety of different complex stable and metastable polymorphs, such as the δ -Ga, γ -Ga, and Ga-II phases^{4,5}. Among them, Ga-II shows thermodynamic stability at high pressure and can be described as interpenetrating body-centered-cubic lattices with each atom having eightfold coordination (Fig. 1c). The structural complexity of liquid and solid Ga makes the study of nucleation and relative equilibria using molecular simulations very challenging.

In order to perform the simulations, two hurdles need to be cleared. One is the long-time scale over which nucleation takes place, whereas the other is an accurate description of the subtle interaction that leads to the complex behavior of Ga. Luckily, much progress has been made in both areas. On the one hand, efficient enhanced sampling methods have allowed studying crystallization in a variety of systems^{13–17}. As far as describing

accurately the Ga interaction is concerned, an ab initio description is called for but running first-principles molecular dynamics (MD) is prohibitively expensive. The solution to this conundrum has been first suggested by Behler and Parrinello¹⁸ and consists in training a neural network (NN) on a large number of appropriately selected set of configurations. This implies calculating accurate total energies and forces, typically at the density functional theory (DFT) level and optimize the parameters of a NN so as to best reproduce these data. The ability of NNs to represent functions of many variables in a continuous way and to interpolate within the training set, allows to obtain a faithful representation of the ab initio potential energies and forces at a much reduced cost. The work of Behler and Parrinello¹⁸ has been developed and applied in many areas. If we confine ourselves to condensed matter applications one can quote refs. ^{19–21} as well as the work of Bartok et al., which uses Gaussian processes rather than NNs to represent the potential²². Here, we choose to use the recently developed deep potential for MD (DeePMD) method^{21,23}. We find that the obtained NN force field can describe the structural and other related properties of α -Ga, β -Ga, Ga-II, and liquid gallium well. Importantly, the phase diagram of gallium in a wide temperature and pressure range are calculated, which shows in good agreement with experimental results. By comparing the nucleation properties of α -Ga and β -Ga, we find that the formation of metastable β -Ga is kinetically favored over the thermodynamically stable α -Ga above 174 K.

Results

NN force field training. A careful choice of the configurations used in the training of the NN^{17,24} is crucial to the success of the NN potential training procedure. As our focus is on the study of the Ga phase diagram and its nucleation process, we need all the characteristic structures of the liquid, the solid, and most importantly, of the nucleation region in which the two phases coexist. Obtaining all the relevant configurations is far from trivial, notably the ones that are far from the equilibrium. To address this issue, we take inspiration from the work by Bonati and Parrinello¹⁷ in which the training configurations were collected from a metadynamics^{25,26} run that used a standard empirical potential to study the liquid–solid phase transition. In the present case, as we set out to study a wide region of the Ga

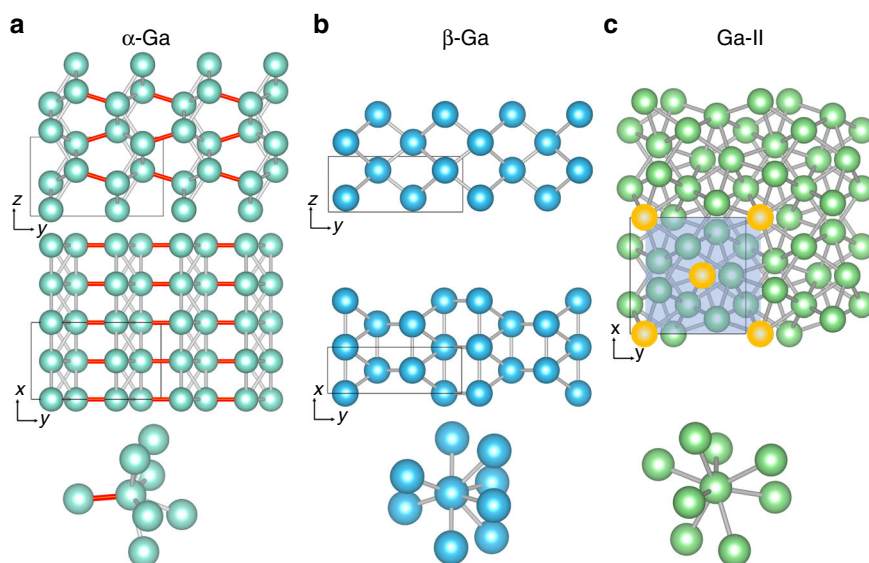


Fig. 1 Crystal structures of solid gallium phases. **a** α -Ga, **b** β -Ga, and **c** Ga-II phases including the atomic nearest neighbor environments of each phase. The covalent bond of the Ga₂ dimer in α -Ga is highlighted with red color. Ga-II can be described as interpenetrated BCC lattices and one of them is highlighted by yellow atoms with a blue-shaded area.

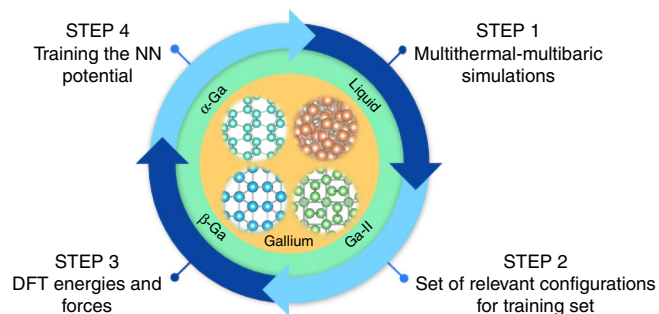


Fig. 2 The neural network training procedure. Four steps, i.e. multithermal-multibaric simulations, set of relevant configurations collection, energies and forces calculation at the DFT level, and the NN potential training, are involved in the iterating procedure.

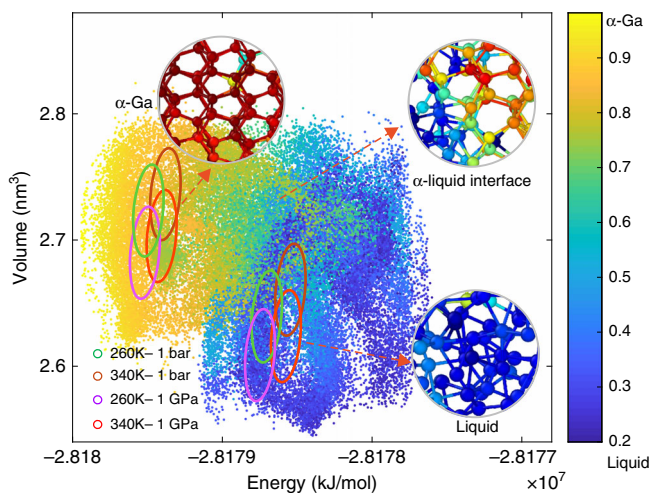


Fig. 3 Energy-volume regions explored with different ensembles. The dots correspond to the region sampled in the multithermal-multibaric ensemble for the nucleation of α -Ga and are colored according to the collective variable defined in Eq. (13). The higher the value of the collective variable, the more solid-like a configuration is. The ellipses show the regions sampled in the isothermal-isobaric ensemble both for α -Ga and liquid at four different conditions, 260 K - 1 bar, 340 K - 1 bar, 260 K - 1 GPa and 340 K - 1 GPa. Inset: structure of α -Ga, α -liquid interface, and liquid gallium.

phase diagram, we generalize the procedure of Bonati and Parrinello¹⁷, and we train our NN on simulations performed in the multithermal-multibaric ensemble^{27,28} so as to present the NN with all the relevant configurations. The NN training procedure includes four steps as illustrated in Fig. 2. These steps are the multithermal-multibaric simulation²⁷, the collection of a set of relevant configurations for the training set, the calculation the energies and forces at the DFT level, and training the NN potential. This procedure is based on the combination of three methods, namely the multithermal-multibaric simulation, the choice of efficient collective variables (CVs)²⁸ and the use of the DeePMD method. By iterating this procedure, we are able to obtain a DFT quality potential.

It is worth noting that the use of multithermal-multibaric simulations is a key feature of our approach, which allows us to build a reliable NN potential and, in turn, to determine the Ga phase diagram. In such simulations, we explore a large variety of arrangements, from those typical of liquid Ga to the α -Ga ones passing through intermediate states. As we report in Fig. 3, these

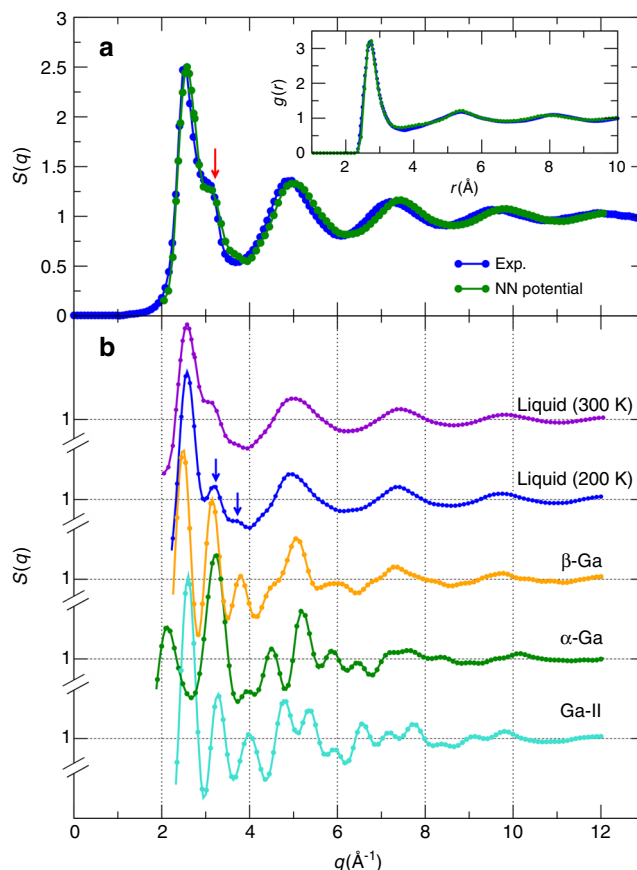


Fig. 4 Analysis of the structure of liquid gallium. **a** Calculated static structure factor $S(q)$ of liquid gallium at 300 K and 1 bar, and comparison with experimental results taken from ref. 2. Inset: radial distribution function $g(r)$ of liquid gallium. **b** Calculated structure factor $S(q)$ of liquid gallium at different temperatures and 1 bar. Structure factors $S(q)$ of α -Ga, β -Ga, and Ga-II are also shown here to highlight the structural similarity between the liquid and β -Ga.

configurations cover a much wider range of energies and volumes compared with standard isothermal-isobaric simulations. Exposing the NN to such a variety of physically relevant configurations is key to training an NN potential valid in a whole-temperature pressure region. Furthermore, employing such simulations allows defining clearly the range of validity where the NN potential can be trusted. This is very important in the context of NNs, as we have no exact way to estimate their ability to generalize.

Structure of liquid gallium. Once the NN potential is obtained, we can use it to study the structure of liquid gallium and its nucleation into different phases. The radial distribution function $g(r)$ and the static structure factor $S(q)$ of liquid gallium obtained with the NN force field are shown in Fig. 4, which are in accordance with the corresponding experimental results. By integrating the $g(r)$ of liquid Ga out to the first minimum, one gets an average coordination number of 11.5, which is similar with liquid metals with close-packed structure, indicating that its atomic environment has a higher degree of isotropy than α -Ga. These features could be expected to contribute to the density anomaly, that results in an expansion of $\sim 3.1\%$ upon freezing into α -Ga. Furthermore, a shoulder on the high- q side of the $S(q)$ is observed, in good agreement with experimental observations. We have also compared the $S(q)$ of the liquid with that of the α , β , and Ga-II phases. The results show that the $S(q)$ of β -Ga exhibits a fair agreement with the positions of the peaks of the liquid

gallium $S(q)$ as shown in Fig. 4b, indicating the underlying structural similarity between β -Ga and the liquid. On the other hand, significant discrepancy could be observed between α -Ga and liquid gallium, in agreement with results reported in previous work⁶. Even though the $S(q)$ of Ga-II also shows some similarity with that of liquid gallium, the role of Ga-II is not important in this case since at ambient temperature and pressure Ga-II is highly unstable⁵. We have also calculated the $g(r)$ and the $S(q)$ of liquid gallium for different temperatures and pressures (see Supplementary Fig. 1). The results show that liquid gallium becomes more structured as the temperature decreases and pressure increases.

Ab initio phase diagram. Now we turn our attention to the solid phases of Ga and their relative equilibrium with liquid gallium. Here, we aim at calculating the phase diagram in the range 240–340 K and 0–2.6 GPa. To achieve this result, we have performed three multithermal–multibaric simulations. These simulations were designed to explore the reversible transitions between the liquid and the different crystal structures (see Supplementary Fig. 2). Then we followed the procedure discussed in ref. 28 to calculate the free energy difference between the liquid and the solid phase $\Delta G_{l \rightarrow s}(T, P)$ as a function of temperature and pressure (see Supplementary Fig. 3). The coexistence lines of the phase diagram could be obtained equating the free energies in the two states $G_{l \rightarrow s}(T, P) = 0$. It should be noted here that the coexistence line between α -Ga and Ga-II is estimated through $G_{l \rightarrow \alpha}(T, P) - G_{l \rightarrow \text{II}}(T, P) = G_{\alpha \rightarrow \text{II}}(T, P) = 0$. The phase diagram of gallium obtained from these three multithermal–multibaric simulations is shown in Fig. 5, and is in agreement with experimental data⁵. The triple point liquid– α -II is located at 280.9 K and 1.58 GPa, which is comparable with the experimental one of 276.2 K and 1.19 GPa⁵.

We have summarized in Table 1 the melting temperatures and lattice parameters of the solid phases investigated here. Both the melting temperatures and the lattice parameters are in good agreement with experiments. For instance, the melting temperature of α -Ga is estimated to be 309.5 ± 6 K, which is slightly higher than the experimental value 302.9 K.

Nucleation of α -Ga and β -Ga. In order to investigate the nucleation behavior of the solid phases, we have performed several metadynamics simulations using the NN potential. Snapshots from the homogeneous nucleation process of α -Ga and β -Ga are illustrated in Fig. 6. Similar features could be observed in both cases. Initially, crystalline embryos form stochastically in the liquid phase. Eventually one cluster dominates and grows steadily into a large crystal. Interestingly, in the y - z plane the crystalline cluster of α -Ga appears relatively spherical, whereas the one of β -Ga is somewhat faceted and shaped as a parallelogram with rounded edges.

From a three-dimensional perspective, the nuclei in both cases are not far from a sphere as can be seen in Fig. 6. This is an important insight into the nucleation mechanism of Ga and will be used below to study this process within the framework of classical nucleation theory (CNT).

Role of Ga₂ dimers in the nucleation of α -Ga. Another question underlying the nucleation process of α -Ga is how the Ga₂ dimers characteristic of this phase evolve. In order to address this issue, we have calculated the free energy surface (FES) as a function of the fraction of solid-like atoms f_s and fraction of Ga₂ dimers f_d at 280 K ($T \approx 0.9 T_m^\alpha$) and 1 bar with 144 Ga atoms in the system. It is worth mentioning that in order to get a converged FES, a much smaller system than the one reported in Fig. 6 is used here. As

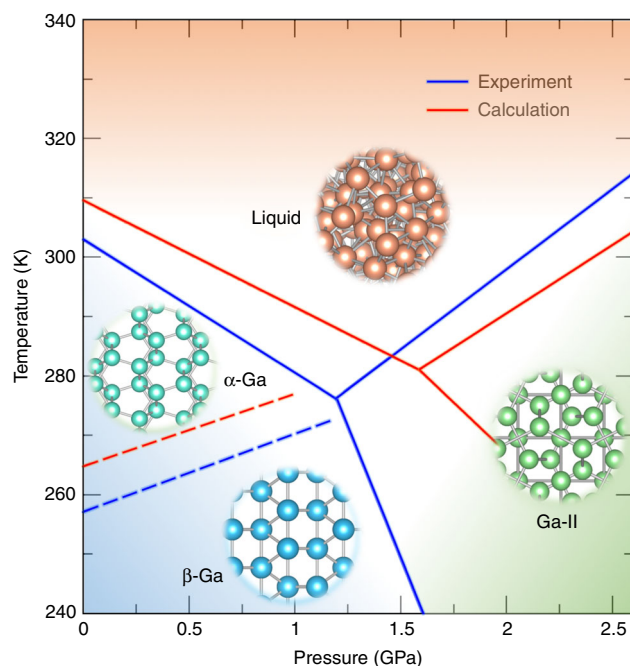


Fig. 5 Phase diagram of gallium. The calculated results from the multithermal–multibaric simulations and the experimental results are colored with red and blue lines, respectively. The coexistence lines between thermodynamically stable phases are shown in solid lines, whereas the coexistence line between metastable β -Ga and the liquid is shown with a dashed line. Inset: Structures of α -Ga, β -Ga, Ga-II, and liquid gallium.

Table 1 Melting temperature and lattice parameters of gallium phases.

Phase	T_m (K)	Lattice parameters (\AA or $^\circ$)	
α	302.9	$a = 4.514$; $b = 7.644$; $c = 4.526$	Exp.
	309.5 ± 6	$a = 4.47 \pm 0.02$; $b = 7.58 \pm 0.04$; $c = 4.48 \pm 0.02$	Cal.
β	256.9	$a = 2.766$; $b = 8.053$; $c = 3.332$; $\beta = 92.03$	Exp.
	264.7 ± 5	$a = 2.69 \pm 0.02$; $b = 7.92 \pm 0.04$; $c = 3.372 \pm 0.02$; $\beta = 92.2 \pm 0.5$	Cal.
Ga-II	314.2	$a = 5.951$	Exp.
	304.3 ± 6	$a = 5.89 \pm 0.03$	Cal.

The calculated (Cal.) and experimental (Exp.) melting temperatures T_m and lattice parameters of the α -Ga, β -Ga and Ga-II phases are listed. The theoretical and experimental data of Ga-II were obtained at 2.6 GPa, whereas the rest of the data were obtained at ambient pressure. The experimental results are taken from ref. 5.

shown in Fig. 7a, the transition between α -Ga and liquid is well described by our CVs. In the solid phase, both the f_s and the f_d are equal to 1, wherein the liquid phase they are approximately equal to 0 and 0.25, respectively. The two basins are roughly linearly connected, indicating the strongly correlation between the Ga₂ dimers and the formation of solid-like atoms. In the liquid state, the Ga₂ dimers are randomly oriented, and the ones along the nucleation direction favors the formation of solid-like atoms. In contrast, the FES as a function of the fraction of solid-like atoms f_s and fraction of Ga₂ dimers f_d of β -Ga nucleation at 240 K ($T \approx 0.9 T_m^\beta$) and 1 bar in the system with 144 atoms is shown in Fig. 7b. Owing to thermal fluctuation, a small fraction of dimers exists also in β -Ga. There is no marked change of f_d during the nucleation process of β -Ga owing to the structure similarity

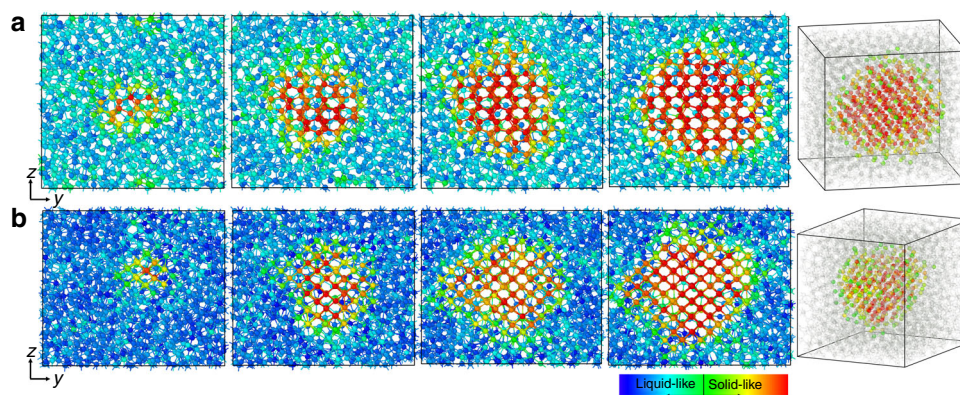


Fig. 6 Snapshots from the nucleation process of gallium. **a** α -Ga at 280 K and 1 bar in a 2650 atoms system and **b** β -Ga at 240 K and 1 bar in a 2400 atoms system. The nuclei of the last snapshots are also shown in three-dimensional perspective and the liquid-like atoms are semi-transparent for clarity. Atoms are colored according to the descriptor defined in Eq. (12).

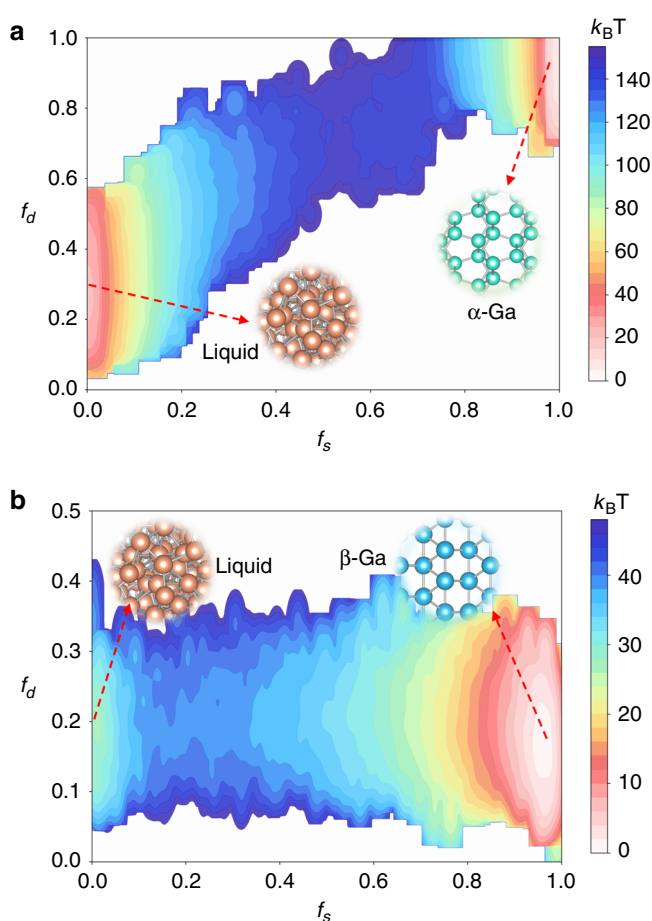


Fig. 7 Reweighted free energy surfaces for the nucleation of gallium. **a** α -Ga at 280 K and 1 bar and **b** β -Ga at 240 K and 1 bar. The free energy surfaces are plotted as a function of the fraction of solid-like atoms f_s and fraction of Ga_2 dimers f_d .

between liquid and β -Ga. These simulations shed light on the evolution of the Ga_2 dimers in the nucleation of α -Ga. However, we could not quantitatively compare the nucleation barriers of α -Ga and β -Ga owing to severe finite size effects.

Competition between α -Ga and β -Ga. In order to overcome the finite size effects and unravel the puzzling competition between

α -Ga and β -Ga in undercooled liquid Ga, we have resorted to the seeding technique²⁹. Under the assumption that CNT holds, this technique can be used to obtain the critical nucleus size and the nucleation energy barrier at moderate supercooling. Usually, a spherical cluster is inserted into the bulk liquid to create the initial simulation system. Here we improved on the standard procedure by using configurations extracted from the metadynamics simulations (Fig. 6) as starting points of the seeding simulations. The advantage of our strategy is that problems such as interface mismatching owing to the insertion of an arbitrary cluster can be avoided, as in our case the cluster is nucleated from the bulk liquid itself. Nevertheless, the configurations extracted from the nucleation of α -Ga and β -Ga were equilibrated for ~ 0.2 ns at the target temperature. We have chosen five initial configurations for each phase, and then we performed MD runs at different temperatures and monitored the cluster size to determine the temperature at which each cluster is critical.

As detailed in the methods section, in the CNT framework the nucleation energy barrier can be calculated as $\Delta G = \frac{1}{2} |\Delta\mu| N_c$, where N_c is the critical nucleus size and $\Delta\mu$ the chemical potential difference between the liquid and solid phase. The latter can be approximated with the enthalpy change at melting as follows: $\Delta\mu = \Delta H_m (1 - \frac{T}{T_m})$ (see ref. 30). The enthalpy of fusion ΔH_m of α -Ga is estimated to be 5.27 kJ/mol, in good agreement with the experimental value, 5.58 kJ/mol³¹.

In Fig. 8, we report the critical nucleus size N_c and the nucleation energy barrier for α -Ga and β -Ga at different temperatures.

First, one can see that the critical size of β -Ga is much larger than that of α -Ga at the same temperature. Furthermore, the curves of the nucleation barrier of the two crystalline phases intersect at a temperature of 174 ± 3 K. Above this temperature, β -Ga has lower nucleation energy barrier than α -Ga, which indicates that undercooling liquid Ga favors nucleation into β -Ga rather than α -Ga. Our results explain experimental findings that the crystallization of liquid Ga with micrometer size or submicrometer size does not form α -Ga but mostly β -Ga^{4,9–11}. We also compare the interfacial free energy γ of these two cases exploiting the fact that $\gamma \propto |\Delta\mu| N_c^{1/3}$ (see Eq. (16)). The results show that the interfacial energy of β -Ga is about three times smaller than that of α -Ga, this could be ascribed to the structural similarity between β -Ga and the liquid we reported earlier.

The nucleation rates J of α -Ga and β -Ga at 170 K and 180 K are estimated following Eq. (18) to be $\sim 2.4 \times 10^{-23}$ and $5.3 \times 10^{-31} \text{ m}^{-3} \text{ s}^{-1}$, respectively. In other words, on average one would have

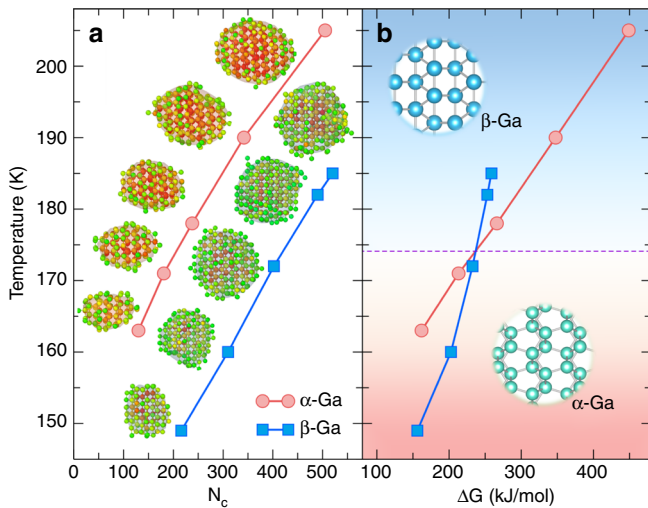


Fig. 8 Nucleation properties of gallium. Temperature dependence of **a** the critical nucleus size N_c and **b** the nucleation energy barrier ΔG of α -Ga and β -Ga. The snapshots in **a** depict the critical nuclei at each temperature studied. The dashed line in **b** separates the figure into two zones, above the line β -Ga has a lower nucleation energy barrier than α -Ga and below the line the opposite is true. The uncertainties of N_c and ΔG are ~ 10 atoms and 8 kJ/mol, respectively.

to wait $\sim 4.8 \times 10^{17}$ and 2.2×10^{25} days to see a single nucleus forming in 1 m^3 of liquid Ga at 170 K and 180 K , respectively. This suggests that homogeneous nucleation will not take place in these conditions and that heterogeneous nucleation is expected to be the normal crystallization mechanism, in accordance with experimental findings¹¹. Our results indicate also why liquid Ga could be experimentally undercooled to extremely low temperatures before the onset of nucleation^{4,9–11}.

Discussion

In this work, we have combined a number of state-of-the-art computational techniques in order to construct a NN force field for gallium, which has many complex bonding and structural properties. The results show that the obtained NN force field can not only describe the structure of liquid gallium well, but it also captures the properties of the investigated solid phases α -Ga, β -Ga, and Ga-II. These properties include the melting temperatures, lattice parameters, and enthalpy of fusion. In addition, we have obtained the phase diagram of gallium in a wide temperature and pressure range, which is in good agreement with experimental findings. Considering the subtleties of the interactions involved, the agreement between theory and experiment is excellent. Furthermore, we have also studied the nucleation of α -Ga and β -Ga, and we show that above 174 K the formation of metastable β -Ga is kinetically favored over the thermodynamically stable α -Ga. The extremely high nucleation barriers for both α -Ga and β -Ga make the crystallization of liquid Ga very hard, which explains the experimental observations of extreme undercooling of liquid Ga. We believe our study thus offers a path to calculate phase diagrams and study the nucleation of complex materials with ab initio accuracy at an affordable cost.

Methods

Multithermal-multibaric simulation. The basic idea behind the variationally enhanced sampling (VES)³² based multithermal-multibaric simulation²⁷ can be more simply grasped if we confine ourselves to a description of the multicanonical ensemble.

We recall that in the VES³² one introduces a functional of the bias $V(\mathbf{s})$,

$$\Omega[V] = \frac{1}{\beta} \log \frac{\int d\mathbf{s} e^{-\beta[F(\mathbf{s})+V(\mathbf{s})]}}{\int d\mathbf{s} e^{-\beta F(\mathbf{s})}} + \int d\mathbf{s} p(\mathbf{s}) V(\mathbf{s}), \quad (1)$$

in which $\beta = 1/(k_B T)$ is the inverse temperature, \mathbf{s} is a set of CVs that are a function of the atomic coordinates \mathbf{R} , the free energy is given within an immaterial constant by $F(\mathbf{s}) = -(1/\beta) \log \int d\mathbf{R} \delta[\mathbf{s} - \mathbf{s}(\mathbf{R})] e^{-\beta U(\mathbf{R})}$, $U(\mathbf{R})$ is the interparticle potential energy, and $p(\mathbf{s})$ is a preassigned target distribution. The minimum of this convex functional is reached for

$$F(\mathbf{s}) = -V(\mathbf{s}) - \frac{1}{\beta} \log p(\mathbf{s}), \quad (2)$$

which amounts to saying that in a system biased by $V(\mathbf{s})$, the distribution is $p(\mathbf{s})$.

In order to generate a multicanonical ensemble, one chooses the potential energy $E = U(\mathbf{R})$ as CV. Using VES we then determine the bias $V(E)$ that leads to a flat distribution in the energy interval $E_1 < E < E_2$. This is obtained by choosing the target distribution

$$p(E) = \begin{cases} 1/(E_2 - E_1), & \text{if } E_1 < E < E_2; \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

By inserting in Eq. (1), the free energy as a function of E becomes

$$F(E) = V(E) - \frac{1}{\beta} \log p(E). \quad (4)$$

On the other hand, as $F(E)$ can be written as

$$F(E) = E - \frac{1}{\beta} \log N(E). \quad (5)$$

one can reconstruct for the chosen interval the density of states $N(E)$. From $N(E)$, the thermal properties can be reconstructed using the fact that the partition function can be written as

$$Z = \int e^{-\beta E} N(E) dE. \quad (6)$$

If the E interval is appropriately chosen, one has in an interval $\beta_1 < \beta < \beta_2$

$$Z \approx \int_{E_1}^{E_2} e^{-\beta E} N(E) dE. \quad (7)$$

In ref. 27, a procedure is described such that once the desired $\beta_1 < \beta < \beta_2$ interval is chosen, the appropriate $E_1 < E < E_2$ interval in Eq. (3) is automatically chosen. Thus with a single simulation one can calculate the system properties in the entire $\beta_1 < \beta < \beta_2$ interval.

Extension to the multithermal-multibaric case is fairly straightforward. In addition to E , one shall also include in the CV set the volume V and perform a VES simulation with the following target distribution:

$$p(E, V) = \begin{cases} 1/\Omega_{E,V}, & \text{if there is at least one } \beta', P' \text{ such that } \beta' F_{\beta', P'}(E, V) < \epsilon \\ & \text{with } \beta_1 > \beta' > \beta_2 \text{ and } P_1 < P' < P_2 \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

where $\Omega_{E,V}$ is a normalization constant, $\beta' F_{\beta', P'}(E, V)$ is the free energy at temperature $T' = 1/(k_B \beta')$ and pressure P' , and ϵ is a predefined energy threshold.

If we want to study first order phase transitions, a CV able to favor the transition between one phase and the other is needed. This leads to a VES problem with three CVs, E , V , and s . Once the run is converged, one can not only compute the properties of the system in a preassigned temperature β_1 – β_2 and pressure P_1 – P_2 region, but also draw the phase boundary. For more details, we refer to ref. 28.

Collective variables. The explicit introduction of an appropriate CV is mandatory to drive the first-order phase transition. We adopt an orientationally targeted order parameter introduced recently by Piaggi and Parrinello²⁸ building on the smooth overlap of atomic positions (SOAP) idea³³. The starting point for the order parameter is the definition of the environment χ around an atom, and the associated local density $\rho_\chi(\mathbf{r})$ is written as

$$\rho_\chi(\mathbf{r}) = \sum_{i \in \chi} e^{-|\mathbf{r}_i - \mathbf{r}|^2 / 2\sigma^2}, \quad (9)$$

where i runs over the neighbors in the environment χ , \mathbf{r}_i are the coordinates of the neighbors relative to the central atom, and σ^2 is the variance of the Gaussian functions. We then measure the difference between the environment χ and χ_0 of the perfect crystal that contains n reference positions $\{\mathbf{r}_j^0\}$ ($j = 1 \dots n$), where $\{\mathbf{r}_j^0\}$ are the crystallographic positions. The two environments are compared by performing the integral,

$$k_{\chi_0}(\chi) = \int d\mathbf{r} \rho_\chi(\mathbf{r}) \rho_{\chi_0}(\mathbf{r}). \quad (10)$$

In the SOAP kernel³³, a spherical average over all the possible orientations of the reference χ_0 is performed. In the work of Piaggi and Parrinello²⁸ the orientation of

χ_0 is instead kept fixed so as to facilitate the nucleation of defect free crystals. Once the rotational symmetry is removed, the integral in Eq. (10) is trivially performed leading to,

$$k_{\chi_0}(\chi) = \sum_{i \in \chi} \sum_{j \in \chi_0} \pi^{3/2} \sigma^3 e^{-|r_i - r_j|^2 / 4\sigma^2}. \quad (11)$$

The kernel function is then normalized

$$\tilde{k}_{\chi_0}(\chi) = \frac{k_{\chi_0}(\chi)}{k_{\chi_0}(\chi_0)} = \frac{1}{n} \sum_{i \in \chi} \sum_{j \in \chi_0} e^{-|r_i - r_j|^2 / 4\sigma^2}. \quad (12)$$

For a system of N atoms, the kernel of each atom i can be denoted by $\tilde{k}_{\chi_0}(\chi_i)$ with $i = 1, \dots, N$. This quantity is a per atom crystallinity measure of the specific phase considered. For this reason, we use it as a fingerprint to follow the nucleation process. The CVs for the whole system is then constructed by counting the number of atoms that satisfy $\tilde{k}_{\chi_0}(\chi_i) > k_0$ where k_0 is chosen so as to separate the solid-like and liquid-like atoms. Finally, the CVs can be written in a continuous and differentiable fashion,

$$s = \sum_i^N \left(1 - \frac{1 - (\tilde{k}_{\chi_0}(\chi_i)/k_0)^p}{1 - (\tilde{k}_{\chi_0}(\chi_i)/k_0)^q} \right), \quad (13)$$

where the parameters p and q control the steepness and the range of the switching function. In this way, $s \approx 0$ for the liquid and $s \approx N$ for the solid structure.

This order parameter can be easily defined for any crystal structure by identifying the reference environments χ_0 . This is a straightforward calculation if the experimental solid phases that we are interested in are known. Furthermore, if the crystal structure of the solid phases is not known one could resort to crystal structure prediction algorithms³⁴.

The kernel that defines the collective variables for α -Ga was defined using $\sigma = 0.4$ and the reference environment χ_0 corresponded to the 13 nearest neighbors of the α -Ga lattice. The kernel for β -Ga was defined using $\sigma = 0.35$ and the reference environment χ_0 corresponded to the 8 nearest neighbors of the β -Ga lattice. The kernel for Ga-II was defined using $\sigma = 0.5$ and the reference environment χ_0 corresponded to the 12 nearest neighbors of the Ga-II lattice.

NN representation of the potential energy surface. Among the many NN-based methods for constructing accurate force fields, we choose here the DeePMD method^{21,23}. As in the work of Bonati-Parrinello¹⁷, our aim is to develop a potential whose validity is restricted to the physical phenomenon of interest namely the reconstruction of a portion of the Ga phase diagram and the nucleation of the relevant crystal phases. In the work of Bonati-Parrinello¹⁷ only $P = 0$ was explored and just one Si crystal structure needed to be considered. This facilitated the construction of the NN potential that was done in two steps. First, the nucleation was simulated using an effective classical potential and relevant configurations were extracted from this run. Second, the relative DFT energies and forces were calculated and on these results the NN potential was trained.

Our aims are more ambitious as we want to cover a broad-phase space region and the Ga iterations are far more subtle than that of Si case dealt within the Bonati-Parrinello work¹⁷. This requires a multi-step procedure. Again we start with a metadynamics calculation²⁵ of α -Ga nucleation at ambient pressure using a classical potential³⁵ and relevant configurations are extracted.

The DFT calculations are performed using the linear density approximations (LDA) exchange–correlation functional. Despite its simplicity, it has been shown that LDA is able to compare well with experiments³⁶ and performs better than more complex functionals like the meta-GGA SCAN³⁷. LDA has limitations, but nonetheless the overall agreement between simulations and experiment is remarkable. However, we note that the coexistence lines of the phase diagram (see Fig. 5) are shifted with respect to the experimental ones. This discrepancy could be owing to the accuracy of the exchange–correlation functional used. For instance, the estimated volume of α -Ga using LDA is $\sim 18.97 \text{ \AA}^3$ per atom, slightly lower than that of experimental value 19.52 \AA^3 . On the contrast, meta-GGA SCAN gives a value of 20.6 \AA^3 . Such discrepancies caused by different functionals could result into different melting temperatures as shown in the literature^{38,39}. Testing other functionals, or levels of theory is beyond the scope of our paper.

A first estimate of the potential NN1 is obtained by training a network on these configurations. The NN1 potential is then used in a number of isothermal–isobaric simulations of α -Ga, β -Ga, and Ga-II nucleation obtaining a new estimated NN2 potential. We then turned to multithermal–multibaric samplings and after two iterations we obtain a good potential able to describe bulk Ga in all the range of temperatures going from 150 K to 340 K and pressures from 1 bar to 2.6 GPa. The root mean square errors of the NN potential on the testing set are equal to 2.8 meV per atom for the energy and 60 meV per \AA for the force.

The use of semiempirical potentials to generate the initial configurations does not limit the generality of our approach. The initial sampling could also have been done with a NN potential trained on ab initio MD simulations of the liquid and the solid phases. The empirical potential, if used, does not need to be particularly accurate, as all the required information of the configurations is then recalculated at the DFT level. Ga is a good example of this scenario since the empirical potential

used here is quite poor and its description of the structure of the liquid phase is far from being adequate. Nevertheless, this did not prevent us from building a DFT quality potential energy surface by following the strategy depicted in Fig. 2.

Nucleation energy barrier and nucleation rate calculations. In the CNT framework, the cost of forming a cluster of the new solid phase from liquid gallium can be expressed as ref. ³⁰:

$$\Delta G^{\text{CNT}} = -\frac{4}{3} \pi R^3 \rho_s \Delta \mu + 4 \pi R^2 \gamma, \quad (14)$$

in which R is the radius of the cluster, ρ_s is the number density of the solid phase, $\Delta \mu$ is the chemical potential difference between the solid and the liquid phase, and γ is the interfacial free energy. The number of atoms in the cluster n can be estimated by $n = \frac{4}{3} \pi R^3 \rho_s$. By inserting this result in Eq. (14) one obtains:

$$\Delta G^{\text{CNT}}(n) = -\Delta \mu n + \sigma n^{2/3}, \quad (15)$$

where $\sigma = (36\pi)^{1/3} \rho_s^{-2/3} \gamma$. ΔG^{CNT} as a function of n exhibits a maximum located at,

$$N_c = \frac{32\pi\gamma^3}{3\rho_s^2 |\Delta \mu|^3}. \quad (16)$$

N_c is the number of atoms in the nucleus or critical cluster. By inserting N_c in Eq. (15) the nucleation energy barrier can be expressed as a function of $\Delta \mu$ and N_c only,

$$\Delta G = \frac{1}{2} |\Delta \mu| N_c. \quad (17)$$

Furthermore, following the approach described in refs. ^{40–42}, one can calculate the nucleation energy barrier J from the expression:

$$J = \rho_l Z f^+ \exp(-\Delta G/(k_B T)), \quad (18)$$

in which $(\rho_l Z f^+)$ is the kinetic prefactor, with ρ_l the number density of the liquid, Z the Zeldovich factor, and f^+ the attachment rate of particles to the critical nucleus. The Zeldovich factor Z in the framework of CNT is

$$Z = \sqrt{|\Delta \mu| / (6\pi k_B T N_c)}, \quad (19)$$

in which k_B refers to the Boltzmann constant. The attachment rate f^+ can be computed as a diffusion coefficient of the cluster size at which the cluster is critical^{40–42}:

$$f^+ = \frac{\langle (N(t) - N_c)^2 \rangle}{2t}, \quad (20)$$

in which $N(t)$ refers to the cluster size at simulation time t . The attachment rates f^+ of α -Ga at 170 K and β -Ga at 180 K are estimated to be $4.7e^{14}$ and $3.8e^{15} \text{ s}^{-1}$, respectively.

MD setup. MD simulations were performed using LAMMPS⁴³ and a development version of PLUMED 2⁴⁴ supplemented by the VES module⁴⁵. The temperature was controlled using the stochastic velocity rescaling thermostat⁴⁶ and the pressure was kept constant employing the isotropic version of the Parrinello–Rahman barostat⁴⁷. The integration time step used is 2 fs and the relaxation times for the thermostat and the barostat were set to 0.1 and 1 ps, respectively.

For the multithermal–multibaric simulations, a bias potential was constructed using VES and employing the energy E , the volume V and s as CVs. Legendre polynomials of order 8 were used in each dimension for a total of 729 variational coefficients. The integrals of the target distribution were performed on a grid of size $40 \times 40 \times 40$. Four multiple walkers were used for the simulations. For α -Ga, the intervals where the polynomials were defined are $-28179800 < E < -28178000 \text{ kJ/mol}$, $2.54 < V < 2.80 \text{ nm}^3$, and $0 < s < 144$. The target distribution was smoothed using Gaussians with $\sigma_E = 200 \text{ kJ/mol}$, $\sigma_V = 0.05 \text{ nm}^3$, and $\sigma_s = 2$. The exploration threshold was set to $\epsilon = 200 k_B T$. The coefficients of the bias potential were optimized every 500 steps using the averaged stochastic gradient descent algorithm with a step size of $\mu = 5 \text{ kJ/mol}$. For β -Ga, the parameters used were the following: $-28179400 < E < -28178100 \text{ kJ/mol}$, $2.48 < V < 2.76 \text{ nm}^3$, $0 < s < 144$, $\sigma_E = 100 \text{ kJ/mol}$, $\sigma_V = 0.05 \text{ nm}^3$, $\sigma_s = 2$, $\epsilon = 40 k_B T$ and $\mu = 1 \text{ kJ/mol}$. For Ga-II the parameter were: $-28179300 < E < -28178300 \text{ kJ/mol}$, $2.42 < V < 2.70 \text{ nm}^3$, $0 < s < 144$, $\sigma_E = 100 \text{ kJ/mol}$, $\sigma_V = 0.05 \text{ nm}^3$, $\sigma_s = 2$, $\epsilon = 40 k_B T$, and $\mu = 1 \text{ kJ/mol}$.

We have used the well-tempered variant of metadynamics (WTMetaD)²⁶ with adaptive Gaussians⁴⁸ to carry out nucleation simulations. The bias factor and initial Gaussian height for α -Ga and β -Ga in the system with 144 atoms were set to be 100 and 50 kJ/mol, and 80 and 20 kJ/mol, respectively. The upper and lower limit of the Gaussian width were set to 2 and 0.2. For the system with 2560 atoms to simulate α -Ga nucleation, a bias factor of 320 and initial Gaussian height 300 kJ/mol were used, whereas for the system with 2400 atoms to simulate β -Ga nucleation, they were set to be 200 and 200 kJ/mol, respectively. The upper and lower limit of the Gaussian width were equal to 1 and 30. The reweighted FES shown in Fig. 7 as a function of the fraction of solid-like atoms f_s and fraction of Ga_2 dimers f_d for the nucleation of α -Ga and β -Ga were calculated following the reweighting procedure given in ref. ⁴⁹.

In the seeding MD simulations, five initial configurations were extracted from the nucleation trajectories of α with 2560 atoms in the system and β -Ga with 2400 atoms obtained by the WTMetaD simulations. After 0.2 ns equilibrium simulation, the sizes of the cluster in the chosen configurations in the case of α -Ga were about 130, 181, 238, 342, and 505 atoms, whereas in the case of β -Ga, they were \sim 216, 310, 402, 490, and 520 atoms.

As the crystal might form in directions that are not aligned with the simulation box during the MD simulation, we introduced the following quantity to avoid such undesired phenomenon as in ref. ²⁸,

$$s_c = \frac{Q_6 - Q_6^l}{Q_6^* - Q_6^l} - \frac{s - s^l}{s^* - s^l}, \quad (21)$$

where Q_6 is the global Steinhardt parameter as defined in refs. ^{50,51} and s is the collective variable that induces the nucleation. As discussed in ref. ²⁸, the rationale behind s_c is that s_c is close to zero only if Q_6 and s increase simultaneously. By keeping s_c close to zero, crystals with orientations different from the desired one are avoided. To achieve this, a restraining potential is applied:

$$V(s_c) = \begin{cases} k(s_c - s_c^0)^2, & \text{if } s_c > s_c^0; \\ 0, & \text{otherwise,} \end{cases} \quad (22)$$

in which $k = 10^5$ kJ/mol. For α -Ga, β -Ga, and Ga-II, s_c^0 equal to 0.25, 0.14, and 0.18, respectively. The distance cutoff in the calculation of Q_6 was 0.35 nm.

DFT calculations setup. DFT⁵² calculations are performed using the Quickstep module⁵³ of the CP2K program⁵⁴. The $3d^{10}3s^2p^1$ electrons were treated explicitly, using LDA as exchange–correlation functional. The single-point energy and forces calculations for the training set used a energy cutoff of 600 Ry. The threshold for energy convergence is set to 10^{-11} and the one related to SCF cycles is set to 10^{-6} . A Fermi Dirac smearing with 0.025 eV of the occupation number at the Fermi level is used, and mixing of the electronic density in k -space is also used. There were 144 atoms in the study of liquid- α and liquid Ga-II phase transitions, where 160 atoms were used in the case of β -Ga. A k -points grid of $2 \times 2 \times 2$ is adopted.

DeePMD setup. The DeePMD-kit package²³ is used for the training of the NN potential and to interface the NN potential to LAMMPS. A NN with five hidden layers is used, with a number of neurons per layer equal to (240,120,60,30,10). The total number of training configurations is 28,000. The network is trained with the ADAM optimizer, with an exponentially decaying learning rate from $1e-3$ to $1e-5$. The prefactors of the energy, the force and the virial terms in the loss functions change during the optimization process from 0.2 to 1, from 500 to 1, and from 0.02 to 0.2.

Data availability

All the data and the input files necessary to reproduce the results contained in this paper are available in the Materials Cloud repository via <https://archive.materialscloud.org/2020.0039/v1>.

Code availability

LAMMPS, Plumed, and DeepMD are free and open source codes available at <https://lammps.sandia.gov>, <https://www.plumed.org>, and <http://www.deepmd.org>, respectively.

Received: 17 October 2019; Accepted: 17 April 2020;

Published online: 27 May 2020

References

- Daeneke, T. et al. Liquid metals: fundamentals and applications in chemistry. *Chem. Soc. Rev.* **47**, 4073–4111 (2018).
- Narten, A. Liquid gallium: comparison of x-ray and neutron-diffraction data. *J. Chem. Phys.* **56**, 1185–1189 (1972).
- Gong, X., Chiarotti, G. L., Parrinello, M. & Tosatti, E. α -gallium: a metallic molecular crystal. *Phys. Rev. B* **43**, 14277 (1991).
- DiCiccio, A. Phase transitions in confined gallium droplets. *Phys. Rev. Lett.* **81**, 2942 (1998).
- Bosio, L. Crystal structures of Ga(II) and Ga(III). *J. Chem. Phys.* **68**, 1221–1223 (1978).
- Bizid, A., Defrain, A., Bellissent, R. & Tourand, G. Neutron diffraction investigation and structural model for liquid gallium from room temperature up to 1 303 K. *J. de Phys.* **39**, 554–560 (1978).
- Gupta, M. K. & O’Sullivan, T. P. Recent applications of gallium and gallium halides as reagents in organic synthesis. *RSC Adv.* **3**, 25498–25522 (2013).
- Heine, V. Crystal structure of gallium metal. *J. Phys. C: Solid State Phys.* **1**, 222 (1968).
- Bosio, L., Defrain, A., Curien, H. & Rimsky, A. Structure cristalline du gallium β . *Acta Crystallogr. Sect. B: Struct. Crystallogr. Cryst. Chem.* **25**, 995–995 (1969).
- Bosio, L. & Windsor, C. Observation of a metastability limit in liquid gallium. *Phys. Rev. Lett.* **35**, 1652 (1975).
- Borisov, B. et al. Peculiarities of gallium crystallization in confined geometry. *Phys. Solid State* **46**, 2286–2291 (2004).
- Parravicini, G. et al. Extreme undercooling (down to 90 K) of liquid metal nanoparticles. *Appl. Phys. Lett.* **89**, 033123 (2006).
- Piaggi, P. M., Valsson, O. & Parrinello, M. Enhancing entropy and enthalpy fluctuations to drive crystallization in atomistic simulations. *Phys. Rev. Lett.* **119**, 015701 (2017).
- Niu, H., Piaggi, P. M., Invernizzi, M. & Parrinello, M. Molecular dynamics simulations of liquid silica crystallization. *Proc. Natl. Acad. Sci.* **115**, 5348–5352 (2018).
- Niu, H., Yang, Y. I. & Parrinello, M. Temperature dependence of homogeneous nucleation in ice. *Phys. Rev. Lett.* **122**, 245501 (2019).
- Zhang, Y.-Y., Niu, H., Piccini, G., Mendels, D. & Parrinello, M. Improving collective variables: the case of crystallization. *J. Chem. Phys.* **150**, 094509 (2019).
- Bonati, L. & Parrinello, M. Silicon liquid structure and crystal nucleation from ab initio deep metadynamics. *Phys. Rev. Lett.* **121**, 265701 (2018).
- Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
- Li, Z., Kermode, J. R. & De Vita, A. Molecular dynamics with on-the-fly machine learning of quantum-mechanical forces. *Phys. Rev. Lett.* **114**, 096405 (2015).
- Schütt, K. T., Sauceda, H. E., Kindermans, P.-J., Tkatchenko, A. & Müller, K.-R. SchNet—a deep learning architecture for molecules and materials. *J. Chem. Phys.* **148**, 241722 (2018).
- Zhang, L., Han, J., Wang, H., Car, R. & Weinan, E. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.* **120**, 143001 (2018).
- Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
- Wang, H., Zhang, L., Han, J. & Weinan, E. DeePMD-kit: a deep learning package for many-body potential energy representation and molecular dynamics. *Computer Phys. Commun.* **228**, 178–184 (2018).
- Behler, J. Perspective: machine learning potentials for atomistic simulations. *J. Chem. Phys.* **145**, 170901 (2016).
- Laio, A. & Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci.* **99**, 12562–12566 (2002).
- Barducci, A., Bussi, G. & Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **100**, 020603 (2008).
- Piaggi, P. M. & Parrinello, M. Multithermal–multibaric molecular simulations from a variational principle. *Phys. Rev. Lett.* **122**, 050601 (2019).
- Piaggi, P. M. & Parrinello, M. Calculation of phase diagrams in the multithermal–multibaric ensemble. *J. Chem. Phys.* **150**, 244119 (2019).
- Espinosa, J. R., Vega, C., Valeriani, C. & Sanz, E. Seeding approach to crystal nucleation. *J. Chem. Phys.* **144**, 034501 (2016).
- Kelton, K.F. Crystal nucleation in liquids and glasses. In *Solid state physics*, vol. 45, 75–177 (Elsevier, 1991).
- Chase Jr, M. W. NIST-JANAF thermochemical tables. *J. Phys. Chem. Ref. Data, Monograph* **9** (1998).
- Valsson, O. & Parrinello, M. Variational approach to enhanced sampling and free energy calculations. *Phys. Rev. Lett.* **113**, 090601 (2014).
- Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).
- Oganov, A. R., Pickard, C. J., Zhu, Q. & Needs, R. J. Structure prediction drives materials discovery. *Nat. Rev. Mater.* **4**, 331–348 (2019).
- Tadmor, E. B., Elliott, R. S., Sethna, J. P., Miller, R. E. & Becker, C. A. The potential of atomistic simulations and the knowledgebase of interatomic models. *JOMom* **63**, 17 (2011).
- Chen, L.-Y., Tang, P.-H. & Wu, T.-M. Analysis of local bond-orientational order for liquid gallium at ambient pressure: two types of cluster structures. *J. Chem. Phys.* **145**, 024506 (2016).
- Remsing, R. C., Sun, J., Waghmare, U. V. & Klein, M. L. Bonding in the metallic molecular solid α -gallium. *Mol. Phys.* **116**, 3372–3379 (2018).
- Zhu, L.-F., Grabowski, B. & Neugebauer, J. Efficient approach to compute melting properties fully from ab initio with application to Cu. *Phys. Rev. B* **96**, 224202 (2017).
- Dorner, F., Sukurma, Z., Dellago, C. & Kresse, G. Melting Si: beyond density functional theory. *Phys. Rev. Lett.* **121**, 195701 (2018).
- Auer, S. & Frenkel, D. Numerical prediction of absolute crystallization rates in hard-sphere colloids. *J. Chem. Phys.* **120**, 3015–3029 (2004).

41. Auer, S. & Frenkel, D. Prediction of absolute crystal-nucleation rate in hard-sphere colloids. *Nature* **409**, 1020 (2001).
42. Espinosa, J., Sanz, E., Valeriani, C. & Vega, C. Homogeneous ice nucleation evaluated for several water models. *J. Chem. Phys.* **141**, 18C529 (2014).
43. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
44. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. Plumed2: new feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).
45. <http://www.ves-code.org>. VES Code, a library that implements enhanced sampling methods based on Variationally Enhanced Sampling written by O. Valsson. For the current version, see <http://www.ves-code.org>.
46. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
47. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
48. Branduardi, D., Bussi, G. & Parrinello, M. Metadynamics with adaptive gaussians. *J. Chem. Theory Comput.* **8**, 2247–2254 (2012).
49. Bonomi, M., Barducci, A. & Parrinello, M. Reconstructing the equilibrium boltzmann distribution from well-tempered metadynamics. *J. Comput. Chem.* **30**, 1615–1621 (2009).
50. Van Duijneveldt, J. & Frenkel, D. Computer simulation study of free energy barriers in crystal nucleation. *J. Chem. Phys.* **96**, 4655–4668 (1992).
51. Reintjes, P., RuizMontero, M. J. & Frenkel, D. Numerical calculation of the rate of crystal nucleation in a lennardjones system at moderate undercooling. *J. Chem. Phys.* **104**, 9932–9947 (1996).
52. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133 (1965).
53. VandeVondele, J. et al. Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach. *Comput. Phys. Commun.* **167**, 103–128 (2005).
54. Hutter, J., Iannuzzi, M., Schiffmann, F. & VandeVondele, J. CP2K: atomistic simulations of condensed matter systems. *Wiley Interdiscip. Rev.* **4**, 15–25 (2014).

Acknowledgements

The authors would like to thank Dr. Tarak Karmakar for carefully reading the manuscript. This research was supported by the NCCR MARVEL, funded by the Swiss National Science Foundation, and the European Union Grant No. ERC-2014-AdG-670227/VARMET. H.N. would like to thank the support by the Research Fund of the State Key Laboratory of Solidification Processing (NPU), China (Grant No.2019-TS-08). P.M.P. was also supported by an Early Postdoc. Mobility fellowship from the Swiss

National Science Foundation. The calculations were performed in Swiss National Supercomputing Center (CSCS) cluster Piz Daint and International Center for Materials Discovery (ICMD) cluster of NPU.

Author contributions

H.N. and M.P. designed the research. H.N. performed the research and analyzed data with help of L.B., P.M.P., and M.P. H.N. and M.P. wrote the paper and all the authors edited the manuscript before submission.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-16372-9>.

Correspondence and requests for materials should be addressed to H.N. or M.P.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020