npj computational materials

www.nature.com/npjcompumats

ARTICLE OPEN

Prediction of stable Li-Sn compounds: boosting ab initio searches with neural network potentials

Saba Kharabadze[1], Aidan Thorn[1], Ekaterina A. Koulakova[1] and Aleksey N. Kolmogorov[1]✉

The Li-Sn binary system has been the focus of extensive research because it features Li-rich alloys with potential applications as battery anodes. Our present re-examination of the binary system with a combination of machine learning and ab initio methods has allowed us to screen a vast configuration space and uncover a number of overlooked thermodynamically stable alloys. At ambient pressure, our evolutionary searches identified an additional stable Li_3Sn phase with a large BCC-based hR48 structure and a possible high-*T* $LiSn_4$ ground state. By building a simple model for the observed and predicted Li-Sn BCC alloys we constructed an even larger viable hR75 structure at an exotic 19:6 stoichiometry. At 20 GPa, low-symmetry 11:2, 5:1, and 9:2 phases found with our global searches destabilize previously proposed phases with high Li content. The findings showcase the appreciable promise machine-learning interatomic potentials hold for accelerating ab initio prediction of complex materials.

npj Computational Materials (2022)8:136 ; https://doi.org/10.1038/s41524-022-00825-4

INTRODUCTION

Accurate density functional theory (DFT) approximations introduced over the past three decades have proven to be powerful practical solutions for quantum mechanical description of materials properties[1–3]. Reliable prediction of structures' thermodynamic stability[4–6] has had a particularly transformative impact on the materials discovery process, as ab initio ground-state searches have been used to screen large chemical spaces and identify new synthesizable materials[4,7]. Despite the growing number of confirmed predictions[8–10], the pace of DFT-guided discovery is ultimately limited by the high cost of ab initio calculations.

Machine learning potentials (MLPs) have emerged as an attractive alternative approach for describing interatomic interactions[11,12]. Being flexible classical models, they can fit DFT potential energy surfaces (PESs) within the typical systematic errors of the reference method but operate at a fraction of the DFT computational cost[9,13–15]. These capabilities have expanded the scope of atomistic simulations well beyond what has been possible with ab initio methods. For example, extended MLP-based molecular dynamics (MD) has been employed for systems with 10^5 atoms[16]. MLP simulations have been used to examine phase transitions[17–23], probe complex surface/interface geometries[12,24,25], simulate crack propagation[26], analyze large-scale defects[27,28], model disordered materials[16,29], and study other complex systems overviewed in ref. [30].

MLP-assisted prediction of ground-state structures differs notably from these materials modeling applications. While ns-long MD trajectories or a-few-nm-sized configurations cannot be checked with DFT directly, pools of candidate structures with lowest Gibbs free energy found in MLP-based searches can and should be examined with the reference method. The benefit of using surrogate models should then be measured against the net cost of their parametrization and the following DFT analysis of viable candidates. MLPs have proven to be essential for finding stable nanoparticle configurations because direct global optimizations with the DFT become impractical for cluster sizes above a few dozen atoms[31–33], while empirical potentials generally lack the accuracy needed to effectively guide ab initio searches. In fact, our recent large-scale comparative study showed a significant advantage of a typical neural network (NN) interatomic model over traditional potentials and led to the revision of several ab initio ground states for Au clusters with 30–80 atoms[34].

The advantage of developing and using MLPs in the exploration of crystalline solids is less apparent because ordered phases tend to have relatively small unit cells. To the best of our knowledge, MLP-guided global structure searches for crystalline ground states have been performed only in a handful of studies[17,35–45] dedicated primarily to unary materials[17,37–45]. This work has resulted in very few predictions of new thermodynamically stable phases, such as Co-Nb-V, Al-Ni, and Al-Ni-Ti[35] and several Mg-Ca alloys[36], as investigations of elemental boron[37–40], carbon[39,41,42], silicon[17,43], and phosphorous[44,45] have either reproduced previously known structures or found metastable ones.

In this study, we aim to illustrate the capabilities of our recently developed framework for MLP-accelerated structure prediction[9,46] by performing a large-scale exploration of Li-Sn alloys. The motivation for (re)investigating this common binary system is twofold. First, the existence of several Li-rich compounds with high electrical conductivity and better ductility compared to other group-XIV elements make them appealing candidate materials for Li-ion battery anodes[47–49]. Second, the binary system has been recently investigated with the state-of-the-art ab initio prediction methods in two systematic studies and shown to host new thermodynamically stable compounds at ambient and elevated pressures[47,50]. The development and application of an accurate NN model for Li-Sn have allowed us to probe over a million structures across the full composition range at different (*P*,*T*) conditions and identify several additional ground states. Particularly surprising are thermodynamically stable BCC-based Li_3Sn and $Li_{19}Sn_6$ phases with large prototypes that appear not to have been observed in any other binary alloys. These findings illustrate the appeal of employing MLPs to accelerate ab initio prediction of complex materials.

[1]Department of Physics, Applied Physics and Astronomy, Binghamton University, State University of New York, PO Box 6000, Binghamton, NY 13902-6000, USA.
✉email: kolmogorov@binghamton.edu

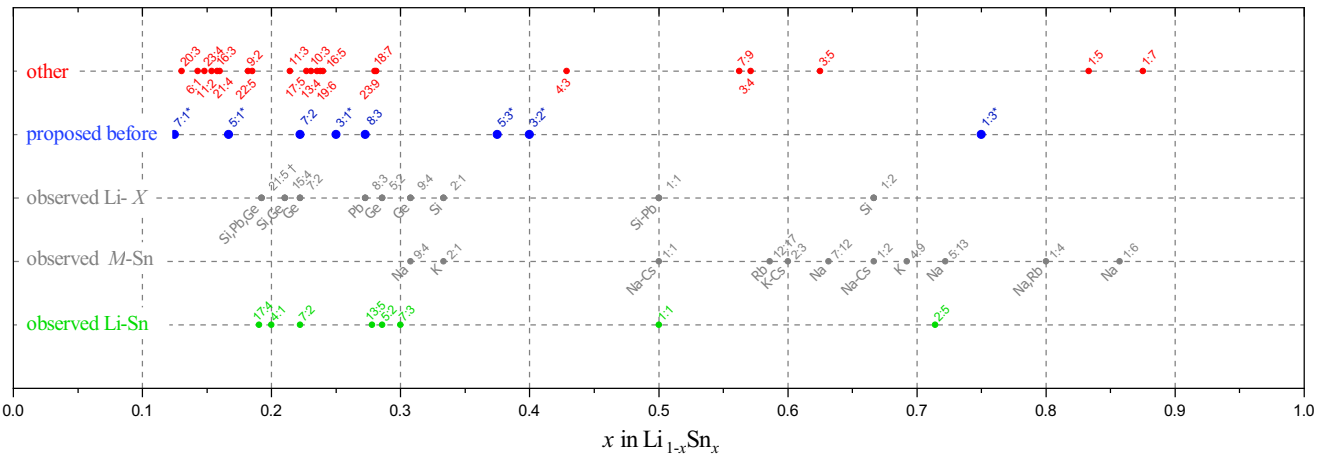


Fig. 1 Relevant compositions in the Li-Sn system explored with our NN-based evolutionary searches. The bottom green set denotes experimentally observed Li-Sn compounds[116]. The two gray groups specify stoichiometries observed in related *M*-Sn and Li-*X* binaries[116]. The blue set refers to thermodynamically stable compounds predicted in previous studies[47,50] for 0 GPa and 20 GPa (the latter are marked with asterisks). The top red group corresponds to additional compositions considered in this study: all possible *m*:*n* ratios with $(m+n) \le 8$ and a few larger subsets with $m = 6$–23 and $n = 1$–9 near regions with known stable compounds.

RESULTS

The main outcomes of the study presented in this Section include the development of our NN-based strategy for finding ab initio crystalline ground states at zero and elevated temperatures, identification of additional stable phases in the Li-Sn system, and analysis of the Li-Sn alloy properties. Further details about the simulation packages, calculation settings, and NN model performance are given in the Methods Section.

Structure prediction strategy

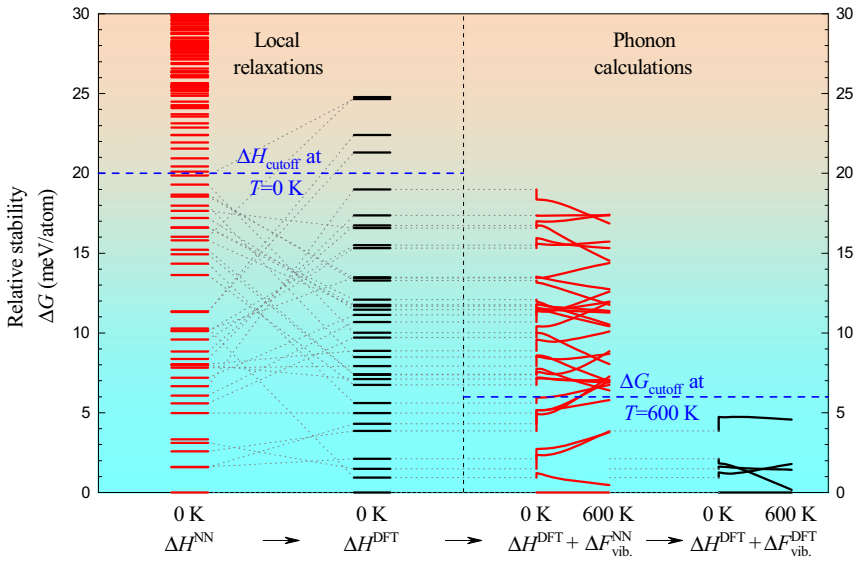
Overview. Identification of ab initio Li-Sn ground states at different (P, T) conditions included evolutionary searches with the NN model at $T = 0$ K, re-examination of viable $T = 0$ K candidates with the DFT, and investigation of phase stability at elevated temperatures at the NN and DFT levels. The highest considered pressure of 20 GPa brackets typical compressions attainable in multi-anvil cells while the highest considered temperatures of 600–800 K are near the melting points of known Li-Sn compounds[51–53]. The global search for stable Li-Sn alloys at $T = 0$ K and $P = 0$ GPa or $P = 20$ GPa was driven by an evolutionary algorithm implemented in MAISE[9]. Detailed in our previous studies[9,54,55], this global optimization engine has been used primarily in combination with DFT to predict or solve complex ground states. It has led to the discovery of several confirmed new prototypes with large unit cells: oP10-FeB_4[54,56], mP20-MnB_4[4], and tI56-CaB_6[55] found without any structural input as well as mP24-Cu_2IrO_3 and tP56-$Na_3Ir_3O_8$ found by seeding the searches with known structures[57,58]. Noteworthy features of the present NN-based evolutionary searches and post-search analysis are described below.

Selection of compositions. A thermodynamically stable $A_{(1-x)}B_x$ phase must have the lowest Gibbs free energy at composition x, negative Gibbs free energy of formation $G_{\text{form}} = G_{A_{1-x}B_x} - (1-x)G_A - xG_B$ with respect to the elemental ground states, and negative relative Gibbs free energies $\Delta G = G_{A_{1-x}B_x} - (x_k - x)/(x_k - x_j)G_k - (x - x_j)/(x_k - x_j)G_j$ with respect to any combination of binary phases j and k. Hence, exhaustive screening of relevant stoichiometries is a requisite for identification of stable compounds. It has been demonstrated[59,60] that unsupervised variable-composition evolutionary searches can result in an efficient location of stable stoichiometries. One of the recent studies on Li-Sn[47] employed such an algorithm and refined solutions with follow-up evolutionary searches at select compositions. Given the richness of the Li-Sn phase diagram and the complexity of the known ground states at non-trivial compositions, we relied on a supervised selection of stoichiometries for our fixed-composition searches.

Figure 1 summarizes available information on Li-Sn and related chemical systems helpful for determining compositions pertinent to this study. The three lowest sets list previously synthesized Li-Sn, *M*-Sn (M = Na-Cs), and Li-*X* (X = Si, Ge, and Pb) alloys. Among the *M*-Sn binaries, Li-Sn stands out as a system with multiple Li-rich compounds and unique stable compositions across the full range. It is evident that the size of the alkali metals is a dominant factor defining stable *M*:Sn ratios, which renders the observed Na-Cs tin alloy compositions not particularly relevant for Li-Sn under ambient conditions. The Li-*X* set, on the other hand, provides important clues regarding possible compositions and morphologies that could be observed in the Li-Sn binary. As discussed below, some of the previously predicted compounds displayed near the top of Fig. 1 were shown to be stable in these prototypes.

Our main focus was on compositions with previously reported analogs in Fig. 1 containing fewer than 40 atoms in the primitive unit cells. The very large observed cF420-$Li_{17}Sn_4$ prototype with 105 atoms in the primitive unit cell was examined directly with the DFT, as finding global minima from scratch for crystalline structures above about 40 atoms is challenging even with evolutionary searches based on surrogate models. We also considered all possible *m*:*n* ratios with $m + n \le 8$ and included a few larger *m*:*n* subsets with $m = 6$–23 and $n = 1$–9 to sample the Li-rich end of the phase diagram (the top row in Fig. 1).

Post-search DFT analysis at $T = 0$ *K*. Upon completion of a NN-based evolutionary run, a subset of viable candidates from all generated local minima was selected for further re-examination with the DFT. The global DFT minimum is captured provided that it is (i) at least a local minimum on the NN PES; (ii) visited by the search algorithm; and (iii) included in the analysis pool[34]. The main adjustable parameter in the selection process is the enthalpy window. By default, we collected all distinct minima within 20 meV/atom ($\approx 2\sigma_E^{\text{Li-Sn}}$) above the NN ground state to account for the NN typical errors in the evaluation of relative enthalpies but reduced the window at some compositions with a large number of competing low-symmetry states. We relied on a 0.95



Total	NN	DFT	NN	DFT
Structures	1,160,000	4,450	2,300	233
CPU hours	952,000	134,000	306	708,000

Fig. 2 An overview of the NN-guided identification of zero-*T* and high-*T* ab initio ground states. The flowchart illustrates the selection of viable candidates following a representative evolutionary search performed for $Li_{18}Sn_4$ unit cells at 0 GPa. The two sets on the left correspond to the lowest-enthalpy NN minima (in red) and the resulting reoptimized DFT minima (in black). The two subsets on the right show the relative Gibbs free energies in the 0–600 K range with the phonon entropy contributions calculated with the NN model (in red) and the DFT (in black). Every subset is referenced to the structure with the lowest enthalpy in the pool at the corresponding level of theory. The table at the bottom shows the tally of the structures examined and computational resources used in the full exploration of the Li-Sn binary.

radial distribution function (RDF)-based dot product cutoff to exclude similar structures and used a 0.1 tolerance to symmetrize the unit cells[9] before relaxing them with DFT.

Analysis of phase stability at high T. Identification of high-*T* ground states is commonly done by including the vibrational entropy contribution to the Gibbs free energy for a pool of low-enthalpy phases found in global structure searches at $T = 0$ K[2,9]. The relative change in Gibbs free energy can reach a few dozen meV/atom for structures with substantially different phonon density of states (DOS) observed, e.g., in *α*-Sn and *β*-Sn[61–63]. Given the demanding nature of phonon calculations at the DFT level, we used the following protocol to examine high-*T* phase stability. Firstly, we calculated thermal corrections with the NN model ($\Delta F^{NN}_{vib.}$) for all distinct candidate structures in the $T = 0$ K DFT pools within a 20 meV/atom window. Secondly, we constructed a convex hull at 600 K using $G^{DFT+NN} = H^{DFT} + \Delta F^{NN}_{vib.}$ and selected up to 5 structures per pool with Gibbs free energies no further than 10 meV/atom from the tie-line. Finally, we evaluated thermal corrections for these structures with DFT and determined high-*T* ground states by examining convex hulls at all temperatures in the 0–800 K range, in intervals of 10 K. Structures included in the thermodynamic stability analysis were dynamically stable, and any structures which remained dynamically unstable were disregarded if a new structure could not be created from their imaginary frequency eigenmodes.

Overview of global structure search results

The information presented in Figs. 2–4, serves to illustrate the size of the configuration space sampled in this work and the reliability of the NN-based structure prediction approach employed for identification of stable Li-Sn phases.

According to our results, populations in the evolutionary searches typically stopped improving after 210, 910, and 2300 configurations were probed in runs with up to 14 atoms, 15–24 atoms, and over 25 atoms per unit cell, respectively. While the stochastic searches are not guaranteed to converge to the global minimum in every instance, we find that sampling 4400 candidate structures with up to 24 atoms is generally sufficient. We reached this conclusion having observed that (i) evolutionary runs up to this system size consistently reproduced or improved the enthalpy of best minima obtained in searches with a divisible number of formula units at the same composition (e.g., in the Li_2Sn, Li_4Sn_2, Li_8Sn_4, and $Li_{16}Sn_8$ series); and (ii) one-two additional runs with the same settings but different starting populations converged to the same ground state in all such tests (about 10% of all screened compositions). Overall, more than 1.1 million local optimizations with the NN model were performed for binary structures with no initial symmetry and an average of 19.4 atoms per unit cell, which required 952,000 CPU hours (see Fig. 2). An equivalent set of DFT calculations would require an estimated 100 million CPU hours (the increase in the average computational cost was estimated by relaxing 20 structures with 5–30 atoms for the same number of steps in the NN and DFT optimizations). The selection and symmetrization of configurations with low NN enthalpies narrowed the pool down to 4450 structures (about 13 structures per pool) and cost considerably less to re-optimize with DFT. Importantly, the near-stable structures were described more accurately with the NN model (6.9 meV/atom root mean square error (RMSE) compared to $\sigma_E^{Li-Sn} = 10.2$ meV/atom) and underwent relatively small geometrical changes in the following DFT local relaxations (2.0 meV/atom enthalpy decrease on average). As a result, the best DFT minima ranked among the most favored ones within the NN pools (the structures with the lowest NN enthalpy remained favored at the DFT level in 45% of all considered cases) and had either matching or lower enthalpies

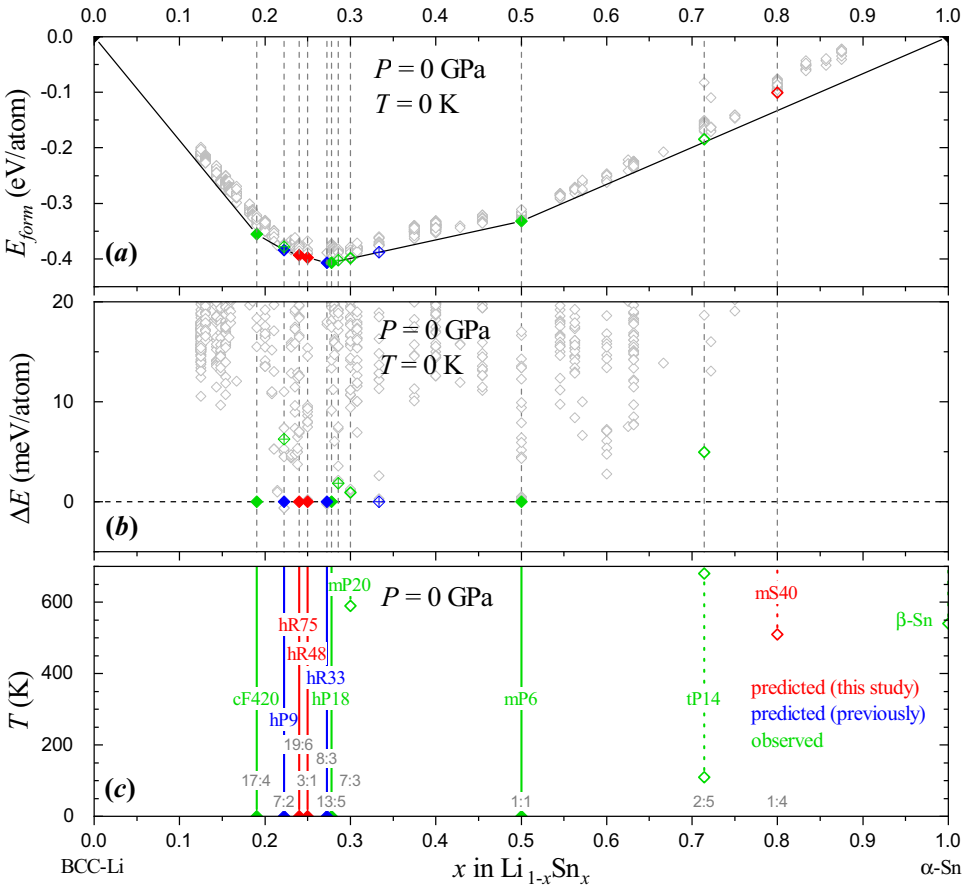


Fig. 3 DFT thermodynamic stability analysis of Li-Sn alloys at $P = 0$ GPa. The Li-B phases observed experimentally, proposed previously[47,50], and predicted in this study are shown in green, blue, and red, respectively. The solid (open) symbols denote phases stable at zero (elevated) temperatures, while the crossed symbols correspond to phases that remain metastable at all considered temperatures. **a** The convex hull at $T = 0$ K. **b** The distance from the convex hull at $T = 0$ K. **c** The temperature range of phase stability. The identified thermodynamically stable phases shown in red are hR75-$Li_{19}Sn_6$, hR48-Li_3Sn, and mS40-$LiSn_4$.

compared to all previously reported Li-Sn phases at $T = 0$ K. It should be noted that repeated runs with randomly initialized populations can produce different pools and miss a local NN minimum corresponding to the global DFT minimum. In this study, all previously observed or proposed Li-Sn ground states with up to 24 atoms appeared as (meta)stable minima in the first and all additional runs.

Our analysis of phase stability at elevated temperatures benefited similarly from the use of the NN model. Following the selection scheme outlined above, we performed phonon calculations for 2300 and 233 structures at the NN and DFT levels, respectively. With the free energy contributions averaging 157 meV/atom at 600 K, the correspondence between the two methods was 7.2 meV/atom. With the RMSE of 1.5 meV/atom and the 4.9 meV/atom average dispersion in *relative* stability at $T = 600$ K (see Supplementary Figs. 1 and 2 for more details), the NN-based screening allowed us to examine an order of magnitude more candidate structures and uncover possible high-T ground states at 0 GPa and 20 GPa.

It is important to note that while DFT systematic errors in calculations of thermodynamic stability may reach a few dozen meV/atom in certain cases[54,64,65], relative energies evaluated for similar structures tend to be far more accurate due to cancellation of errors[36,66]. Our tests performed for different functionals and convergence settings (Supplementary Tables 1 and 2) demonstrate that the relative enthalpies can be resolved to within a fraction of 1 meV/atom for most considered structures with underlying BCC morphologies. Moreover, while Gibbs free energies of formation at room temperature calculated with DFT (Supplementary Fig. 3) and fitted to experimental data (in ref. [52]) differ by up to 60 meV/atom because the Li-Sn phases are referenced to unrelated elemental ground states, the DFT convex hull is demonstrated below to be reliable for determining stable Li-Sn phases. These findings are in line with the conclusions of two large-scale studies reporting a 97% agreement between experiment and DFT predictions of phase stability in 80 binary alloys that tend to adopt similar close-packed configurations[67] and a lower 84% agreement in 41 metal-boron binaries that feature diverse boron frameworks[4]. Temperature estimates for transitions between near-degenerate phases with small differences in free energy corrections can have significant uncertainties of a few hundred Kelvin[36,68].

Our main findings regarding competing Li-Sn phases at 0 GPa and 20 GPa are summarized in Figs. 3 and 4. The DFT results allowed us to refine known $T = 0$ K convex hulls (top), identify metastable alloys (middle), and establish temperature ranges of phase stability (bottom). Namely, two additional ambient-pressure ground states are introduced next to the previously proposed ones[47,50] and several revised high-pressure ground states substantially redefine the Li-rich end of the convex hull predicted in ref. [47]. The collection of local minima obtained in our NN-based evolutionary searches illustrates the presence of many more

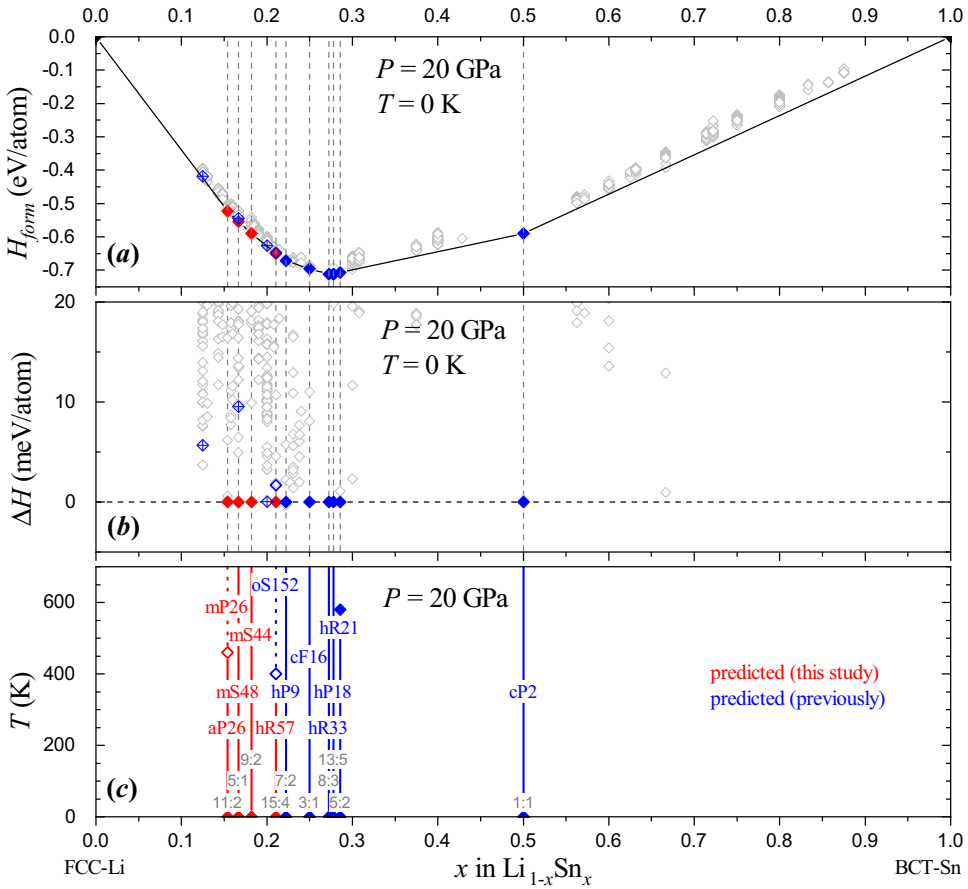


Fig. 4 The same as in Fig. 3, but at $P = 20$ GPa. **a** The convex hull at $T = 0$ K. **b** The distance from the convex hull at $T = 0$ K. **c** The temperature range of phase stability. The four putative ground states in red identified with our NN-based evolutionary searches redefine the Li-rich end of the previously proposed $T = 0$ K convex hull.

phases within the 20 meV/atom window than found with the DFT-based random sampling approach[50]. The inclusion of thermal corrections indicates that only a few considered metastable phases have a chance of becoming thermodynamically stable.

Stability and morphology of Li-Sn phases

The variety of observed Sn alloy morphologies[69–72] can be attributed to the particular position of the element in the periodic table. Similarly to the C-Ge group-XIV members, Sn displays a propensity to form covalent bonds and crystallizes in the open diamond structure (α-Sn) at low (P, T). In contrast to the light isoelectronic elements, Sn has significantly reduced bond integrals essential for the s–p hybrid orbital favorability[73] and, consequently, it adopts various ground states with metallic bonding, such as β-Sn above 286 K[61] or closed BCT structure between 11 and 32 GPa[69,71]. Li is an archetypical sp metal with the BCC ground state under ambient conditions transforming into FCC above ~ 5 GPa[74,75].

In combination, Li and Sn have been observed to form BCC-based alloys at high Li concentrations and compounds with covalent frameworks at high Sn concentrations. The prevalence of BCC motifs can be viewed as a consequence of the negative chemical pressure induced by the larger Sn ($a^{Sn}_{BCC} = 3.81$Å) that further favors the BCC-Li lattice ($a^{Li}_{BCC} = 3.43$Å) over the high-pressure FCC-Li polymorph. The largest magnitude of the formation enthalpy achieved around the 4:1 composition (Figs. 3 and 4) is consistent with the octet rule[47,76–78].

At ambient pressure, $Li_{17}Sn_4$ is the most Li-rich synthesized alloy in this binary system. Goward et al. performed a comprehensive XRD analysis of $Li_{22}X_5$ (X = Ge, Sn, and Pb) compounds and determined that the slight compositional variability across the series originates from selected occupation of Li sites[79]. The ordered cF420 ($F\bar{4}3m$) structural model used to simulate $Li_{17}Sn_4$[47,50] has only one Sn Wyckoff site (16e) with the full (8+6)-atom BCC coordination that produces tetrahedral agglomerates of Sn polyhedra (see Fig. 5a). The other three Sn Wyckoff sites, 16e, 24g, and 24f, are surrounded by 13 Li atoms and generate isolated Sn atoms sufficient to fill up all space with interlocking polyhedra. As in previous studies[47,50], the phase is found to be thermodynamically stable in our DFT calculations.

The observed oS36-Li_7Sn_2 phase represents a regular BCC lattice with fully occupied sites. The Sn atoms in this Li_7Ge_2 prototype are not fully segregated, making one of the two Sn Wyckoff sites have a Sn neighbor in the second shell at 3.20 Å (Fig. 5b). Interestingly, previous ab initio studies revealed a lower-energy decoration of the BCC lattice at this composition[47,50]. The simpler hexagonal hP9 phase with the Li_7Pb_2 prototype does not contain directly interacting Sn atoms and ends up being ~ 6 meV/atom more stable in (semi)local DFT approximations refs. [47,50,80] and Supplementary Table 1. Even though the energy difference reduces to ~ 3 meV/atom at the strongly-constrained and appropriately-normed (SCAN) level (Supplementary Table 1), the oS36 phase appears to be only metastable under typical synthesis conditions. We reproduce the previously reported finding[47] that the inclusion of vibrational entropy changes the relative free

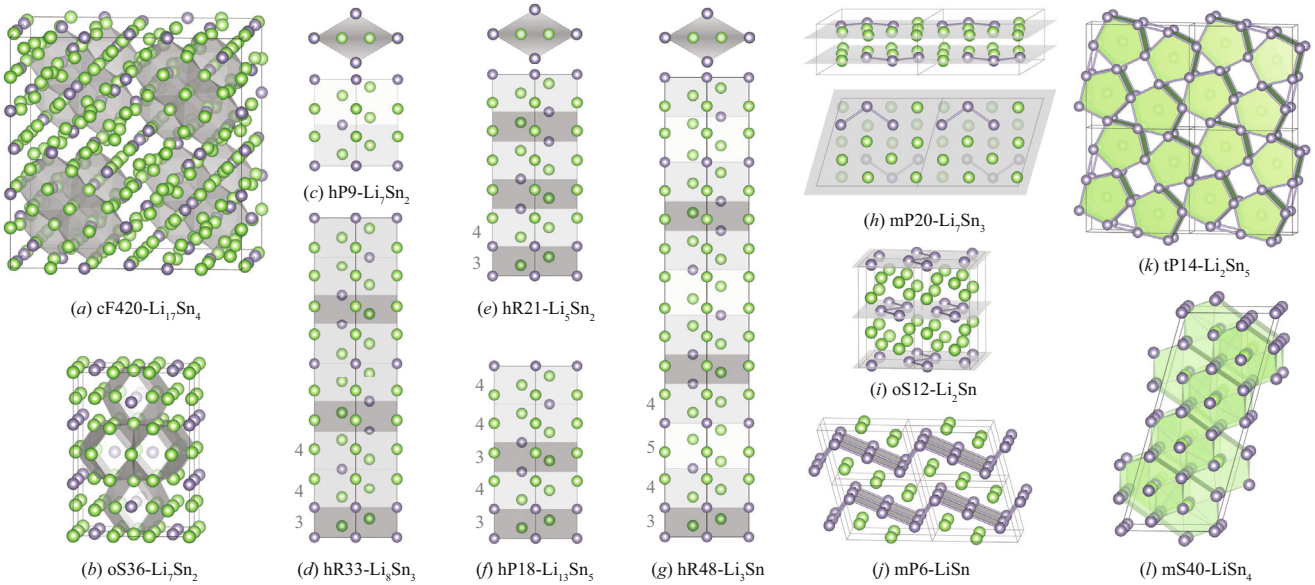


Fig. 5 Select crystal structures of Li-Sn alloys (near)stable under ambient conditions. The displayed phases have been observed experimentally (**a**, **b**, **e**, **f**, **h**, **l**, **k**), proposed previously in refs. [47,50] (**c**, **d**), and predicted in this study (**g**, **l**). The Li and Sn atoms are shown in green and gray, respectively, while the shaded polyhedra, sections, and planes highlight distinctive morphological features discussed in the text. The [111] BCC alloys shown in **c**–**g**, along with our predicted hR75-$Li_{19}Sn_6$ not included due to its large size, can be uniquely defined with sequences of Sn blocks, e.g., |3454| for hR48-Li_3Sn. These images were generated with the VESTA package[117].

energy insignificantly (by less than 1 meV/atom up to 800 K in Supplementary Fig. 4), which is not unexpected given the similarity of the oS36 and hP9 morphologies.

The stable hR33-Li_8Sn_3 phase[47,50] with the Li_8Pb_3 prototype has been proposed[47] to explain the nominal compound synthesized in 1996 by Gasior et al.[81]. Just as hP9-Li_7Sn_2, it exhibits an ordered pattern along the [111] crystallographic direction of the BCC lattice and can be represented with a hexagonal unit cell (Fig. 5d). In contrast to hP9-Li_7Sn_2, it has a peculiar distribution of the Sn atoms: as discussed in refs. [47,50], the occupation of neighboring sites creates Sn-Sn dimers with an unusually short 2.90 Å interatomic distance.

The known hP18-$Li_{13}Sn_5$ and hR21-Li_5Sn_2 phases at nearby compositions bear a strong resemblance to hR33 and contain Sn-Sn dimers of length 2.91 Å (on two of the three Sn Wyckoff sites) and 2.92 Å (on the only 6*c* Sn Wyckoff site), respectively. Our Perdue-Burke-Ernzerhof (PBE) calculations agree with the previous results[47] regarding the thermodynamic stability of the former and the metastability of the latter. As summarized in Supplementary Table 1, hR21-Li_5Sn_2 is 1.8 meV/atom above the hP18-$Li_{13}Sn_5 \leftrightarrow$ mP6-LiSn tie-line in this approximation but marginally breaks the convex hull by −0.4 meV/atom in the SCAN calculations. The higher sensitivity of the relative stability to the DFT flavor, in this case, is likely caused by the fairly different morphology of the reference mP6-LiSn phase discussed below.

In order to systematize the description of such BCC structures featuring the same hexagonal base, we introduce a convenient notation that builds on the detailed structural analysis presented in ref. [47]. As can be seen from Fig. 5c–g, Sn atoms 3, 4, and 5 layers apart along the *c* axis are shifted by 0, 1/3, and 2/3 unit cell fractions within the *a*-*b* base, respectively. Therefore, Sn sequences can be used to uniquely specify these trigonal crystal system structures and easily deduce whether they belong to the hexagonal or rhombohedral lattice system. Indeed, sequences with the total number of Sn layer vertical separations divisible by three (e.g., |45| with 2 Sn atoms in the 9-atom primitive unit cell) result in no net lateral shift and, hence, correspond to hexagonal structures. All other sequences (e.g., |34|) must be tripled to ensure orthogonality of the *c* axis with the base and, thus, define rhombohedral structures.

At the 3:1 composition, only metastable ambient-pressure phases have been proposed so far[47,50] but our evolutionary searches for 12:4 unit cells uncovered another thermodynamically stable member of the BCC family, hR48-Li_3Sn with the |3435| sequence. The rhombohedral structure with a 16-atom unit cell appeared in both NN evolutionary runs after 849, and 1491 local optimizations, although it placed 17 meV/atom above a cF16 BCC phase and ranked 4th in the NN pool. In the considered DFT approximations, it was found to be the lowest-energy 3:1 configuration and stable with respect to the decomposition into the previously proposed hP9-Li_7Sn_2 and hR33-Li_8Sn_3 BCC-based ground states by about −2 meV/atom at $T = 0$ K (Supplementary Table 1). Following phonon calculations indicated that hR48-Li_3Sn should remain stable at elevated temperatures.

The finding motivated us to examine BCC structures more closely for other possible ground states. The cluster expansion method has been widely used to explore combinatorially large configuration spaces on given lattices[82,83] but an exhaustive screening of all possible decorations of large unit cells would require a separate dedicated study. Fortunately, the observed favorability of the [111] BCC motifs allowed us to focus on the most promising subspace. Instead of relying on traditional atomic clusters, we expanded the total energy as a linear combination of various blocks comprised of 3, 4, or 5 layered units described above:

$$E = \sum_i N_i c_i, \tag{1}$$

where c_i is the energy of block i and N_i is the number of such blocks in a structure. Taking advantage of the phenomenological model's easy interpretability, we considered different blocks, marked with an overbar, to determine stable unit combinations and create competitive BCC structures. We observed that the trivial single-unit $\{i\} = \{\overline{3}, \overline{4}, \overline{5}\}$ set, providing a linear interpolation of energies in the 4:1–2:1 composition range, offered a 8.1 meV/atom accuracy comparable to that of the NN model. The

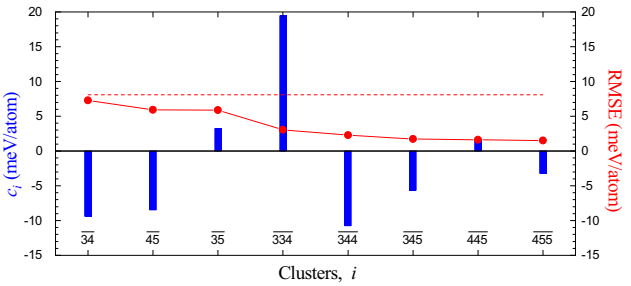


Fig. 6 Strength of the expansion coefficients (left vertical axis) and RMSE (right vertical axis) in the description of [111] BCC phases using a linear model based on blocks with Sn atoms separated by 3, 4, or 5 layers. The |3453|, phase, for example, has $N_{\overline{3}} = N_{\overline{5}} = N_{\overline{345}} = 1$ and $N_{\overline{4}} = N_{\overline{34}} = N_{\overline{45}} = 2$ in Eqn. (1) (symmetric combinations, such as $\overline{34}$ and $\overline{43}$, are grouped together). The RMSE was calculated for a different number of expansion terms from 3 ({$\overline{3}$, $\overline{4}$,$\overline{5}$} giving 8.1 meV/atom marked with the dashed red line) to 11. The displayed coefficient strengths were evaluated for the full set of 11 blocks.

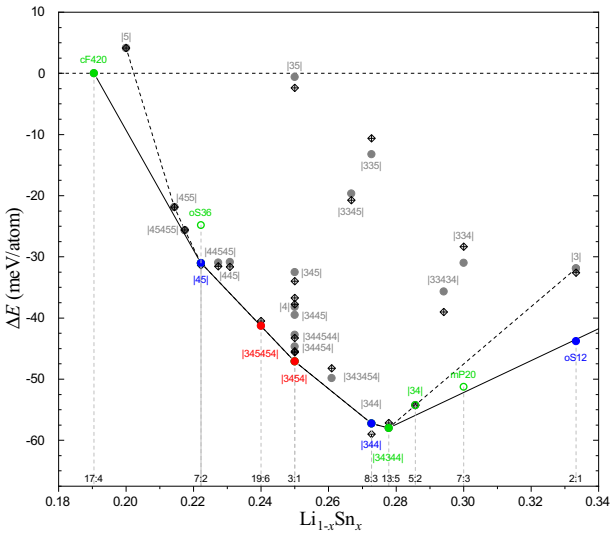


Fig. 7 Stability of Li-rich phases relative to cF420-$Li_{17}Sn_4$ and mP6-LiSn. In addition to the evaluation with the DFT (solid circles), the stability of phases with the [111] BCC morphology, denoted with block sequences explained in the text, is also approximated with a phenomenological model (crossed diamonds). The DFT convex hulls shown with dashed and solid lines correspond to the [111] BCC subset and all considered phases, respectively. The Li-Sn phases observed experimentally, proposed previously, and predicted in this study are shown in green, blue, and red, respectively.

final $\{i\} = \{\overline{3}, \overline{4}, \overline{5}, \overline{34}, \overline{45}, \overline{35}, \overline{334}, \overline{344}, \overline{345}, \overline{445}, \overline{455}\}$ set approximated the energies with an RMSE of 1.5 meV/atom and showed that the $\overline{34}$, $\overline{45}$, $\overline{344}$, and $\overline{345}$ blocks are beneficial while $\overline{35}$ and $\overline{334}$ are detrimental for structure stability (Fig. 6). Based on this information, we constructed a number of possible members with up to 6 units and plotted their DFT relative formation energies (Fig. 7).

This analysis led to identification of yet another thermodynamically stable BCC phase, hR75-$Li_{19}Sn_6$. Comprised of six layered units, the |345454| structure is far more difficult to find with ab initio searches because it has an unusual composition and a 25-atom primitive rhombohedral unit cell with a large $c/a \approx 15$ ratio. Our three NN evolutionary runs with 10^4 local relaxations failed to locate the presumed 19:6 global minimum and converged to a monoclinic mS50 structure 2.1 meV/atom (7.3 meV/atom in the PBE) above hR75. According to our DFT tests (Supplementary Table 1), hR75-$Li_{19}Sn_6$ lies essentially on the hP9-Li_7Sn_2 ↔ hR48-Li_3Sn tie-line and below the hP9-Li_7Sn_2 ↔ hR33-Li_8Sn_3 tie-line by 1–2 meV/atom.

Alloys with higher Sn content ($x \geq 0.3$) adopt stable configurations with distinctive extended Sn fragments or frameworks[50]. The observed mP20-Li_7Sn_3 phase[84] is a peculiar decoration of the BCC lattice resulting in Sn zig-zag trimers with relatively short bonds of 2.99 and 3.00 Å. It has been found slightly metastable by 1–2 meV/atom with respect to hP18-$Li_{13}Sn_5$ and mP6-LiSn in the standard DFT calculations (Supplementary Table 1)[47,50]. Our phonon calculations suggest a stabilizing effect of the vibrational entropy but the flatness of the relative free energy curve makes it difficult to give a reliable estimate of the phase transition temperature (Fig. 3). Moreover, our SCAN calculations show that it may be a true $T = 0$ K ground state.

The previously proposed oS12-Li_2Sn phase[50] was found to be the most stable 2:1 configuration in our calculations as well. The structure contains planar zig-zag Sn chains separated by two Li layers. In contrast to other Li-Sn phases discussed so far, the Li and Sn sites have 12-atom coordinations and are distributed on an HCP lattice. The PBE calculations with an ultrasoft pseudopotential placed it just 1 meV/atom above the hP18-$Li_{13}Sn_5$ ↔ mP6-LiSn tie-line[50]. Our PBE and SCAN results (Supplementary Table 1) indicate that oS12-Li_2Sn may actually be thermodynamically stable at zero and elevated temperatures (Fig. 3).

The experimental mP6-LiSn phase[85] is another example of a stable BCC alloy. A clustered population of the lattice sites creates fully connected zig-zag Sn nets shown in Fig. 5j. Our evolutionary searches for various unit cell sizes of the 1:1 composition located several DFT local minima within a 20 meV/atom energy window (Fig. 3 and Supplementary Fig. 1). The most dramatic entropy-driven stabilization relative to mP6, from 14.9 meV/atom at 0 K down to 6.4 meV/atom at 600 K, occurred for a tI24 phase observed in ref. [86], but none of the considered polymorphs became stable in our calculations (Supplementary Figs. 1 and 4) up to the reported melting temperature of 758 K at this composition[51].

The most Sn-rich synthesized alloy, tP14-Li_2Sn_5 with the Hg_5Mn_2 prototype[87], can be viewed as a network of Li-intercalated pentagonal prisms (Fig. 5k). Located 1.2 meV/atom above the mP6-LiSn ↔ α-Sn tie-line at $T = 0$ K, the phase is predicted to become stable above 110 K (Fig. 3).

Our screening of competitive high-T compounds produced a viable FCC-based mS40-$LiSn_4$ phase (Fig. 5l) that breaks the tP14-Li_2Sn_5 ↔ α-Sn tie-line at 510 K. This temperature estimate is particularly sensitive to the description of pure Sn. The difficulty of reproducing $\alpha \rightarrow \beta$ transition temperatures and pressures for group-XIV Si and Sn has been pointed out in several studies[61–63,88–91]. The overestimated temperature value for the α-Sn to β-Sn transformation obtained in our standard PBE calculations is 540 K. A lower free energy of the reference β-Sn could shift the mS40-$LiSn_4$ stabilization to significantly higher temperatures.

To the best of our knowledge, there have been no attempts yet to synthesize Li-Sn alloys at high pressures, which makes Sen and Johari's ab initio study[47] an important baseline for stability of binary compounds up to 20 GPa. Our evolutionary searches performed at 20 GPa reproduced all the previously reported ground states for compositions above $x = 2/9$. The 7:2, 8:3, 13:5, and 5:2 compounds retain their ambient-pressure BCC prototypes discussed above, while 3:1 and 1:1 compounds adopt well-known simpler cF16 ($D0_3$) and cP2 ($B2$) prototypes, respectively. Compared to the ambient-pressure results, there are very few competing phases within the 20 meV/atom window above $x = 1/3$.

The closest to stability phase is tI12-$LiSn_2$, which is 1 meV/atom above cP2-LiSn ↔ BCT-Sn as seen in Fig. 4.

For compositions below $x = 2/9$, our searches produced several different stable Li-Sn phases that reshape the Li-rich part of the previously proposed convex hull at 20 GPa. Given the relatively low symmetry of these phases and the apparent lack of well-defined underlying motifs, their structural features are better illustrated with the RDF plots shown in Supplementary Fig. 5. It can be seen that Sn atoms in these alloys are separated by at least 4.0 Å, and the Sn local environments consist of 14-17 Li atoms 2.48–3.06 Å away, with more Sn-rich phases generally having shorter Sn-Li interatomic distances. Despite the morphological commonalities, the slight variations lead to noticeable differences in enthalpies.

At the 7:1 stoichiometry, we obtained several competing configurations with low symmetry. The most stable aP32 structure is −1.9 meV/atom below aP16 proposed previously[47]. Both aP16 and aP32 have Sn sites with well-defined shells of 16 Li atoms within the 2.6–3.0 Å range. By linking Sn atoms with 4.5-5.0-Å connections, one can discern different extended Sn frameworks with channels of variable size along *a*, larger in aP32, filled with Li atoms.

According to our global optimization results, the most Li-rich ground state at 20 GPa actually occurs at the 11:2 composition and has a low-symmetry aP26 structure. It makes all 7:1 phases metastable by at least 3.7 meV/atom. While both the NN and DFT methods favor the aP26 polymorph, the latter indicates that another 11:2 phase, mP26, is only 0.4 meV/atom above the putative ground state at $T = 0$ K and becomes stable at 460 K. Notwithstanding the near degeneracy in the full temperature range, the two structures feature different Li coordinations of the two Sn sites: 16 and 16 in aP16 versus 15 and 17 in mP26.

The most significant improvement in stability with respect to the previously proposed ground states is observed at the 5:1 stoichiometry. The best configurations identified in previous[47] and our studies have the same Pearson symbol (mS48) and space group (*C*2/*m*) but the latter is more stable by 10.7 meV/atom; in fact, the former is only metastable by 8.9 meV/atom relative to our updated set of ground states at neighboring compositions (Supplementary Table 2). Morphologically, the previous phase features only 4*i* Li/Sn Wyckoff positions, Sn sites coordinated with 15 and 17 Li atoms 2.61–3.06 Å away, and a Sn framework with Li-filled channels along *b*, whereas ours contains four 8*j* Li Wyckoff positions, Sn sites coordinated with 14 Li atoms only 2.55–2.75 Å away, and a Sn framework without apparent extended pores.

A mS44-Li_9Sn_2 phase found in our study represents another stable compound at a different stoichiometry, being at least −8 meV/atom lower in enthalpy with respect to the best 5:1 and 15:4 phases. Compared to the compressed structures discussed so far, mS44 has Sn-Li interatomic distances dispersed over a wider 2.50–3.36 Å range. The inclusion of this phase in the convex hull destabilizes the previously proposed tI50-Li_4Sn[47] by about 1 meV/atom (see Supplementary Table 2).

Finally, we sampled the unusual 15:4 composition because phases with the cI76-$Cu_{15}Si_4$ prototype ($I\bar{4}3d$) have been observed in related Li-Si and Li-Ge binaries[92,93] and because another oS152-$Li_{15}Si_4$ prototype (*Fdd*2) has been shown to become stable in the Li-Si system above 7.9 GPa[94]. Our DFT calculations at 20 GPa indicate that the cubic structure is unstable by over 60 meV/atom but the orthorhombic one is actually −1.3 meV/atom below the tie-line connecting previously proposed 5:1 and 7:2 ground states[47] and could have been considered stable. We find that oS152-$Li_{15}Sn_4$ breaks the updated 9:2 and 7:2 tie-line as well, albeit by only −0.6 meV/atom. However, our evolutionary searches identified a better hR57 structure −1.8 meV/atom lower in enthalpy compared to oS152. This presumed $T = 0$ K ground state with a large rhombohedral unit cell belongs to the [111] BCC family and can be represented simply as |4555|. Our phonon calculations demonstrate that oS152 does stabilize over hR57 above 400 K.

The impressive number of viable phases with large low-symmetry unit cells appearing at ambient and high pressures testifies to the exceptional complexity of the Li-Sn phase diagram. While we find it encouraging that our extensive screening accelerated with machine learning uncovered several additional phases with lower Gibbs free energies, it cannot be considered exhaustive. Future explorations of the Li-Sn system with higher-accuracy MLPs and DFT approximations may lead to further revision of the phase diagram.

Analysis of electronic and other properties of ambient-pressure Li-Sn compounds

Electronic, mechanical, and electrochemical properties of Li-Sn alloys have been investigated in several ab initio studies[47–50,95]. Analyses of the projected DOS and Bader decomposition data have indicated the expected metallicity of all Li-Sn alloys and transfer of charge from Li to Sn. Interestingly, the charge redistribution in alloys with high Li concentrations exceeding 4:1 has been shown to leave Li atoms in either ionized (+0.8) or nearly neutral (−0.3 to 0.0) states[47]. The manifold of the electronic states near the Fermi level has primarily the Sn-*p* character but displays increasing Li-*s*/*p* contributions in Li-rich compounds[47,49]. It has been pointed out that addition of Li makes the dominant bonding in Li-Sn alloys evolve from covalent Sn-Sn/Li-Sn with a characteristic pseudo-gap to metallic Li-Li[49].

Examination of DOS profiles near the Fermi level can indeed provide insights into materials' various properties ranging from transport, superconductivity, and magnetism to stability. The presence of a pseudo-gap and the population of bonding states has been found to correlate with structure's favorability in numerous cases[5,68,58]. For instance, our studies of related chemical systems have shown that the unusual stoichiometry of LiB_x ($x \sim 0.9$) and the stability of $NaSn_2$ result from the placement of the Fermi level near the bottom of a well-defined DOS valley[5,68,96].

Figure 8 shows that most of the known and proposed Li-Sn alloys, especially those with $x \geq 0.5$, have a significant DOS at the Fermi level and may not conform to this simple model. In fact, it is not easy to disentangle the hybridization and charge transfer factors that determine the shape and the population of the electronic states with the Sn-*p* and Li-*s*/*p* characters. The Sn-*s* states, on the other hand, exhibit little hybridization and provide information about the behavior of Sn-centered orbitals. With the large *s*-*p* splitting disfavoring the formation of the traditional sp^2/sp^3 hybrids in pure Sn open structures, the low-lying *s* states split first and only then does the antibonding *s* set mix partially with the *p* states (Fig. 8o)[97,98].

In the *β*-Sn, mS40-$LiSn_4$, tP14-Li_2Sn_5, and mP6-LiSn phases, the Sn-*s* states disperse into a single band that partially overlaps in energy with the bottom of the Sn-*p* band. As the number of Sn nearest atoms within the 3.08–3.40 Å range changes from 6 to 8–12, 6, and 4 in this series, the Sn-*s* bandwidth gradually decreases. In oS12-Li_2Sn with zig-zag chains, the dispersion reduces down to about 4 eV which creates a 1-eV separation between the Sn-*s* set and the Sn-*p* and Li-*s*/*p* manifold. All shown phases with higher Li content do not have extended Sn frameworks and feature fairly localized Sn-*s* states. The number and the width of the DOS peaks are determined by the distribution of isolated Sn trimers (in mP20-Li_7Sn_3), Sn dimers (for BCC phases with $0.24 < x < 0.277$), and single Sn atoms (in hP9-Li_7Sn_2 and cF420-$Li_{17}Sn_4$). For example, the presence of the 2.9-Å Sn-Sn dimers within the type-3 units in the [111]-BCC phases (Fig. 8e–h) produces two 0.7-eV peaks centered about 2 eV apart, while the orbital overlaps between Sn atoms 4.7 Å apart within the [111]-BCC base or the type-4 blocks give rise to the central 1-eV peaks (Fig. 8c, e–h). The lower edge of the Sn-*p* band moves in

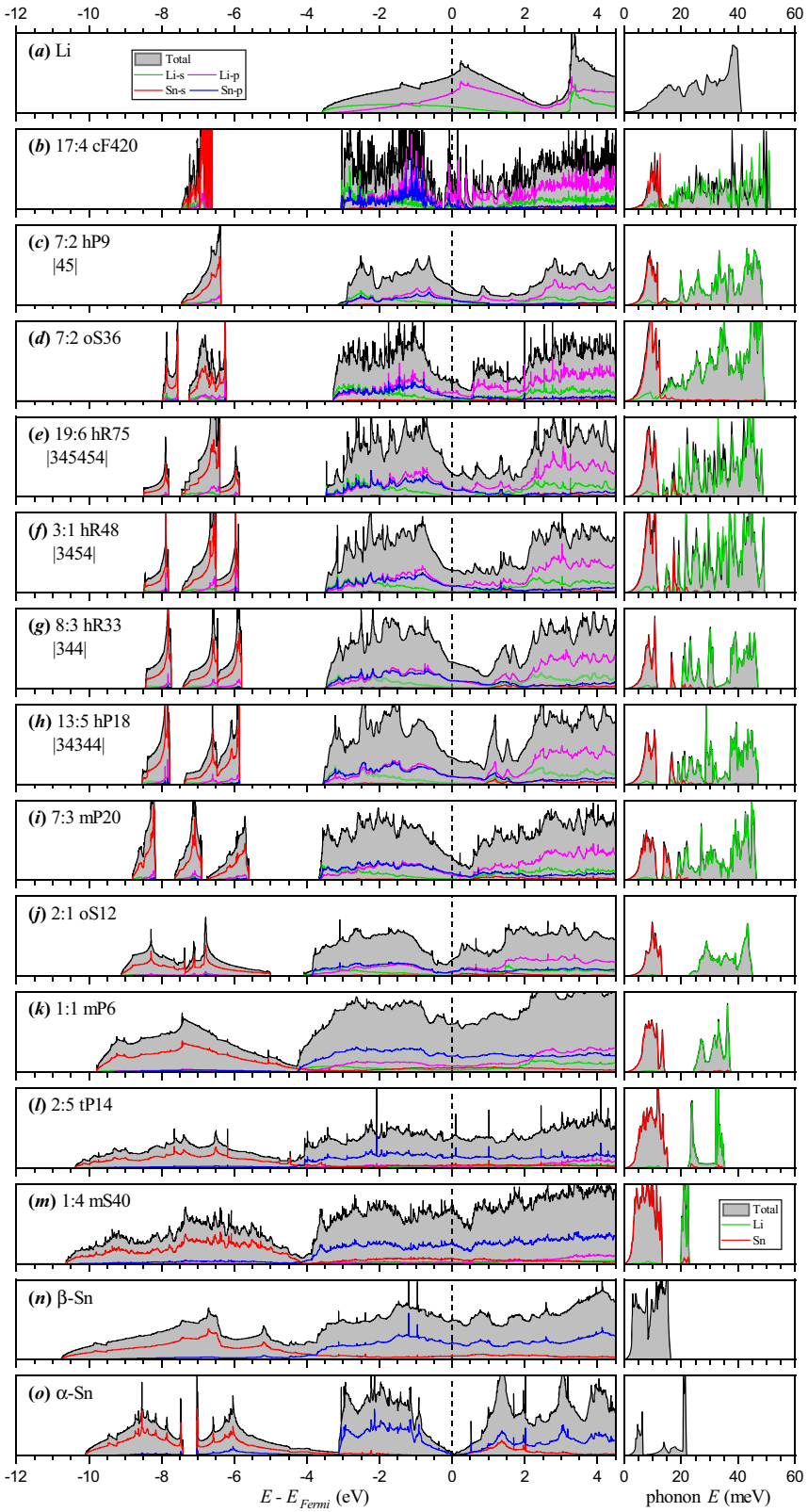


Fig. 8 Electron and phonon DOS calculated with the DFT for known and predicted ambient-pressure Li-Sn phases. The colored lines illustrate the DOS projected onto *s* and *p* orbitals for electrons (left panels) or onto Li and Sn atoms for phonons (right panels).

concert with the bandwidths of the Sn-*s* states, shifting from − 4 eV in oS12-Li_2Sn to − 3 eV in cF420-$Li_{17}Sn_4$. Interestingly, the lowest states in the manifold in oS12-Li_2Sn and hP9-Li_7Sn_2 have the Li-*s*/*p* character due to a larger dispersion of these nearly free electron states that can be seen more clearly on band structure plots for select phases (Supplementary Figs. 6–8).

The change in volume upon Li insertion is one of the critical factors for durability of Li-ion battery anodes. It has been shown

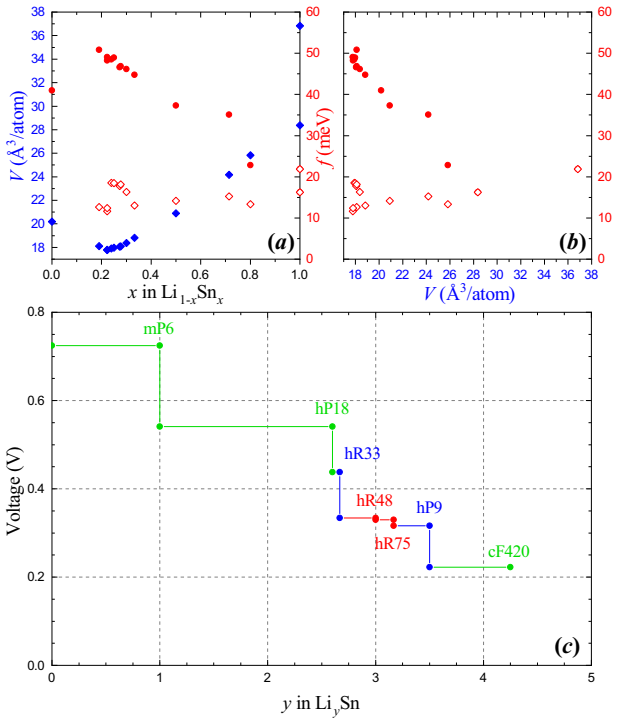


Fig. 9 DFT analysis of select Li-Sn phase properties. a Atomic volume (blue diamonds), highest Li phonon frequencies (solid red circles), and highest Sn phonon frequencies (empty red diamonds) evaluated at ambient pressure as a function of the alloy composition. **b** The same highest Li and Sn phonon frequencies as a function of the atomic volume. **c** Voltage for stable Li_ySn alloys refined through addition of two hR48-Li_3Sn and hR75-$Li_{19}Sn_6$ ground states predicted in this study.

that the calculated atomic volume in Li-Sn alloys notably deviates from the linear dependence on the Li content[47,48]. Our results in Fig. 9 demonstrate that the atomic volume correlates with the formation energies in Fig. 3 across the $Li_{1-x}Sn_x$ composition range. The lowest values occur in compounds with the maximum transfer of charge from Li to Sn near the 4:1 stoichiometry.

A similar trend in the full composition range can be observed for vibrational properties using the phonon DOS results in Fig. 8. The highest phonon frequencies of 41 meV in BCC-Li, 22 meV in α-Sn, and 16 meV in β-Sn indicate that the force constants are higher in Sn than in Li because the mass factor would scale down the Sn frequencies by a larger $\sqrt{m_{Sn}/m_{Li}} \approx 4.1$. The top edge of the Sn DOS peak downshifts to 13 meV in FCC-based mS40-$LiSn_4$, goes back up to 15 meV in tP14-Li_2Sn_5 based on a 3D Sn framework, and eventually settles at 12 meV in BCC-based Li-rich alloys. The highest frequencies of the Li modes actually increase with the first addition of Sn, reaching a maximum of 51 meV in cF420-$Li_{17}Sn_4$, and then decrease with the Sn concentration, dropping to 23 meV in $LiSn_4$. Plotted versus the atomic volume in Fig. 9b, the highest phonon frequencies in all Li-Sn phases exhibit a clearer trend: the BCC-Li and α-Sn outliers fall closer to the linear fit, which leaves tP14-Li_2Sn_5 with hard Li modes as the only significant deviation.

Finally, we comment on how the predicted ambient-pressure phases could be detected. Electrochemical cycling measurements have been used to identify Li-Sn alloys in several studies[99–101]. We update the previously reported average voltage plot[50] by including the hR75-$Li_{19}Sn_6$ and hR48-Li_3Sn phases predicted in our study. Since these phases deepen the convex hull by only 1–2 meV/atom, the additional features near $y = 3$ in Li_ySn (Fig. 9) are very fine and would be difficult to observe. However, their powder X-ray diffraction (XRD) patterns differ considerably from those in previously observed or proposed BCC phases and would provide strong evidence of their formation (Supplementary Fig. 9). The combination of electrochemical and XRD measurements has been recently used to detect a new $NaSn_2$ phase[102] predicted in our earlier ab initio study[68].

DISCUSSION

Our re-examination of the Li-Sn binary illustrates that well-known systems may still host additional synthesizable compounds and that computationally demanding ab initio prediction of thermodynamically stable materials can be effectively boosted with MLPs. Given the surprising scarcity of successful MLP-assisted predictions of stable compounds so far, our study was focused on devising guidelines for constructing reliable models, identifying viable structures, and establishing stability trends that increase the chance of finding ab initio ground states. To this extent, the combination of the evolutionary sampling and stratified training allowed us to build a practical Li-Sn NN potential with a relatively modest 10.2 meV/atom accuracy but a reliable performance in unconstrained searches. The model construction involved less than 2% of the total computational cost required for the systematic exploration of the binary system. Our screening of over 1.1 million (2.3 thousand) crystalline phases at zero (elevated) temperature with the NN potential along with the following DFT analysis of select candidates cost only an estimated 0.1–1% of the traditional ab initio searches. As a result, we discovered overlooked thermodynamically stable hR48-Li_3Sn and mS40-$LiSn_4$ at ambient pressure and several possible ground states at 20 GPa. In contrast to the previously identified stable hP9-Li_7Sn_2 and hR33-Li_8Sn_3 phases[47,50] that have known Li_7Ge_2 and Li_8Pb_3 prototypes in related binaries and can be naturally identified by chemical substitution, our large hR48 and mS40 structures appear to have no analogs in other chemical systems. Moreover, after introducing a phenomenological model for the [111] BCC family of Li-rich alloys, we constructed another stable hR75-$Li_{19}Sn_6$ phase with an even larger unit cell that could not be found with our standard evolutionary searches. These findings highlight the benefit of combining different prediction strategies that utilize available experimental knowledge, explore configuration spaces with global optimization algorithms, and involve rational design based on stability analysis. Since the thermodynamic stability of the additional compounds predicted with this method is ultimately determined by the accuracy of the employed DFT approximation, we performed additional tests with various DFT flavors. Our results indicate that the relatively small formation enthalpy differences of a few meV/atom are fairly insensitive to the systematic and numerical errors due to the similarity of the competing configurations. For the same reason, the vibrational entropy contribution to the Gibbs free energy was found to be an insignificant factor for differentiating Li-rich phases with related BCC motifs. Consequently, the vast majority of detected metastable phases in the 20 meV/atom enthalpy window, a common criterion for viable candidates, proved to be irrelevant in the Li-Sn system. The proposed thermodynamically stable phases are expected to be synthesizable and detectable with standard powder XRD measurements. The formation of hR48-Li_3Sn and hR75-$Li_{19}Sn_6$ would lead to further refinement of the electrochemical profile at the Li-rich end of the complex phase diagram.

METHODS

Density functional theory calculations

All DFT calculations were performed with the Vienna ab-initio simulation package[103–106]. To ensure proper description of materials under high pressures, we chose projector augmented wave potentials[107] with semi-core electrons of Li (1*s*) and Sn (4*d*). Unless specified otherwise, we used

the PBE exchange-correlation functional[108] within the generalized gradient approximation (GGA)[109] and the energy cutoff of 500 eV. Select phases were examined with a higher 700 eV energy cutoff and within the local density approximation[110,111] or SCAN in the meta-GGA[112]. All crystalline structures were evaluated with dense ($\Delta k \sim 0.02$ Å^{-1}) Monkhorst-Pack k-point meshes[113].

Neural network interatomic potential

A NN model of the Behler-Parrinello type was constructed using an automated iterative scheme implemented in our MAISE-NET framework[9] that features an evolutionary sampling algorithm to generate representative reference structures and a stratified training protocol to build multicomponent models on top of elemental ones[46]. Resulting datasets consist mostly of partially optimized crystal structures with up to 10 atoms obtained in short evolutionary runs at 0, 10, 30, and 50 GPa and contain a small fraction (10–20%) of compressed and expanded configurations of basic prototypes and found minima. About 10% of structures randomly selected from the former group are set aside for testing. For elemental Li and Sn, we relied on our previously developed NN models[9] with 51-10-10-1 architectures and $\sigma_E^{\mathrm{Li}} = 2.3$ meV/atom and $\sigma_E^{\mathrm{Sn}} = 8.3$ meV/atom testing RMSEs. The fitting was done on 4717 energy and 32,550 force data points for Li and 4346 and 29,694 data points for Sn. The binary Li-Sn NN with a 145-10-10-1 architecture with 1880 adjustable parameters (the total of $(8_{\mathrm{LiSn}} + 43_{\mathrm{LiLiSn}} + 43_{\mathrm{LiSnSn}}) \times 10_{\mathrm{Li}} + (8_{\mathrm{SnLi}} + 43_{\mathrm{SnSnLi}} + 43_{\mathrm{SnLiLi}}) \times 10_{\mathrm{Sn}}$ inter-species weights connecting the 8 pair and 43 triplet input symmetry functions with 10 neurons in the first hidden layer of each elemental NN) was fitted to 6046 energy and 46,410 force data in binary structures with up to 32 atoms and compositions in the $0.125 \le x \le 0.875$ range generated in four MAISE-NET cycles. The optimization of the binary model for 120,000 steps resulted in $\sigma_E^{\mathrm{Li-Sn}} = 10.2$ meV/atom and $\sigma_F^{\mathrm{Li-Sn}} = 49.0$ meV/Å RMSEs over a test set of 671 structures. The distribution of the NN energy errors is displayed in Supplementary Fig. 10. All elemental and binary models were based on our standard set of Behler-Parrinello symmetry functions with a 7.5 Å cutoff radius[34]. In total, the extension of the elemental Li and Sn models to the binary one cost under 35,000 CPU hours on compute nodes with CPU model Intel Xeon Gold 5218@2.30 GHz.

Evolutionary structure searches

Global structure optimizations were performed with our MAISE package[9]. At each selected composition, we ran separate evolutionary runs with different numbers of formula units. Typically, structures had between 1 and 8 formula units and did not exceed 40 atoms in the primitive unit cell. Randomly generated populations of 32 members were evolved for up to 300 generations with standard evolutionary operations. Namely, 8 new members were constructed from a single parent structure through mutation (random atom displacements, atom swaps, and unit cell distortions) while 24 offspring were created from two parents through crossover (combination of two roughly equal parts obtained with planar cuts)[54]. Child structures were locally relaxed with the NN potentials for up to 300 Broyden-Fletcher-Goldfarb-Shanno minimization steps and assigned a fitness based on the final enthalpy. Our fingerprint method based on the RDF[9,55] was used to identify similar structures and decrease their survival probability.

Vibrational property analysis

Both NN and DFT phonon calculations were performed in the harmonic approximation with the finite displacement method as implemented in Phonopy[114]. We used symmetry-preserving expansions of primitive or conventional unit cells to generate supercells with 72–216 atoms and applied 0.1 Å displacements. The Gibbs free energy corrections due to the vibrational entropy were included via summation over $20 \times 20 \times 20$ grids in the Brillouin zone approximating the integral $\Delta F_{\mathrm{vib.}} = k_B T \int_0^\infty d\omega g(\omega) ln[2sinh(\hbar\omega)/2k_B T]$.

DATA AVAILABILITY

The Li-Sn NN model, ID 031C0A79, can be downloaded at https://github.com/maise-guide/maise/. Relevant Li-Sn structures are given in the supplementary material. Other data supporting the findings of this study is available from the corresponding author upon request.

CODE AVAILABILITY

The MAISE and MAISE-NET codes are freely available for download at https://github.com/maise-guide/.

Received: 28 February 2022; Accepted: 3 June 2022;
Published online: 28 June 2022

REFERENCES

1. Hautier, G., Jain, A. & Ong, S. P. From the computer to the laboratory: materials discovery and design using first-principles calculations. *J. Mater. Sci.* **47**, 7317–7340 (2012).
2. Oganov, A. R., Pickard, C. J., Zhu, Q. & Needs, R. J. Structure prediction drives materials discovery. *Nat. Rev. Mater.* **4**, 331–348 (2019).
3. Alberi, K. et al. The 2019 materials by design roadmap. *J. Phys. D: Appl. Phys.* **52**, 013001 (2018).
4. Van Der Geest, A. & Kolmogorov, A. Stability of 41 metal–boron systems at 0 GPa and 30 GPa from first principles. *Calphad* **46**, 184–204 (2014).
5. Kolmogorov, A. N. & Curtarolo, S. Theoretical study of metal borides stability. *Phys. Rev. B* **74**, 224507 (2006).
6. Sun, W. et al. The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2**, e1600225 (2016).
7. Jain, A. et al. The Materials Project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
8. Hautier, G., Ong, S. P., Jain, A., Moore, C. J. & Ceder, G. Accuracy of density functional theory in predicting formation energies of ternary oxides from binary oxides and its implication on phase stability. *Phys. Rev. B* **85**, 155208 (2012).
9. Hajinazar, S., Thorn, A., Sandoval, E. D., Kharabadze, S. & Kolmogorov, A. N. MAISE: Construction of neural network interatomic models and evolutionary structure optimization. *Comput. Phys. Commun.* **259**, 107679 (2021).
10. Lepeshkin, S., Baturin, V., Uspenskii, Y. A. & Oganov, A. R. Method for simultaneous prediction of atomic structure and stability of nanoclusters in a wide area of compositions. *J. Phys. Chem. Lett.* **10**, 102–106 (2018).
11. Behler, J. Perspective: machine learning potentials for atomistic simulations. *J. Chem. Phys.* **145**, 170901 (2016).
12. Bartók, A. P. et al. Machine learning unifies the modeling of materials and molecules. *Sci. Adv.* **3**, e1701816 (2017).
13. Zuo, Y. et al. Performance and cost assessment of machine learning interatomic potentials. *J. Phys. Chem. A* **124**, 731–745 (2020).
14. Behler, J. First principles neural network potentials for reactive simulations of large molecular and condensed systems. *Angew. Chem. Int. Ed.* **56**, 12828–12840 (2017).
15. Deringer, V. L., Caro, M. A. & Csányi, G. Machine learning interatomic potentials as emerging tools for materials science. *Adv. Mater.* **31**, 1902765 (2019).
16. Deringer, V. L. et al. Origins of structural and electronic transitions in disordered silicon. *Nature* **589**, 59–64 (2021).
17. Behler, J., Martoňák, R., Donadio, D. & Parrinello, M. Metadynamics simulations of the high-pressure phases of silicon employing a high-dimensional neural network potential. *Phys. Rev. Lett.* **100**, 1–4 (2008).
18. Khaliullin, R. Z., Eshet, H., Kühne, T. D., Behler, J. & Parrinello, M. Graphite-diamond phase coexistence study employing a neural-network mapping of the ab initio potential energy surface. *Phys. Rev. B—Condens. Matter Mater. Phys.* **81**, 18–21 (2010).
19. Eshet, H., Khaliullin, R. Z., Kühne, T. D., Behler, J. & Parrinello, M. Microscopic origins of the anomalous melting behavior of sodium under high pressure. *Phys. Rev. Lett.* **108**, 1–4 (2012).
20. Sosso, G. C., Miceli, G., Caravati, S., Behler, J. & Bernasconi, M. Neural network interatomic potential for the phase change material GeTe. *Phys. Rev. B - Condens. Matter Mater. Phys.* **85**, 1–13 (2012).
21. Baldock, R. J., Pártay, L. B., Bartók, A. P., Payne, M. C. & Csányi, G. Determining pressure-temperature phase diagrams of materials. *Phys. Rev. B* **93**, 1–9 (2016).
22. Mocanu, F. C. et al. Modeling the phase-change memory material, $Ge_2Sb_2Te_5$, with a machine-learned interatomic potential. *J. Phys. Chem. B* **122**, 8998–9006 (2018).
23. Kruglov, I. A., Yanilkin, A., Oganov, A. R. & Korotaev, P. Phase diagram of uranium from ab initio calculations and machine learning. *Phys. Rev. B* **100**, 1–7 (2019).
24. Artrith, N., Hiller, B. & Behler, J. Neural network potentials for metals and oxides—first applications to copper clusters at zinc oxide. *Phys. Status Solidi (B) Basic Res.* **250**, 1191–1203 (2013).
25. Boes, J. R. & Kitchin, J. R. Modeling segregation on AuPd(111) surfaces with density functional theory and Monte Carlo simulations. *J. Phys. Chem. C.* **121**, 3479–3487 (2017).

26. Shen, L., Wang, Y. & Lai, W. Development of a machine learning potential for the study of crack propagation in titanium. *Int. J. Press. Vessels Pip.* **194**, 104514 (2021).
27. Stricker, M., Yin, B., Mak, E. & Curtin, W. Machine learning for metallurgy II. A neural-network potential for magnesium. *Phys. Rev. Mater.* **4**, 103602 (2020).
28. Goryaeva, A. M. et al. Efficient and transferable machine learning potentials for the simulation of crystal defects in bcc Fe and W. *Phys. Rev. Mater.* **5**, 103803 (2021).
29. Qian, X., Peng, S., Li, X., Wei, Y. & Yang, R. Thermal conductivity modeling using machine learning potentials: application to crystalline and amorphous silicon. *Mater. Today Phys.* **10**, 100140 (2019).
30. Hart, G. L., Mueller, T., Toher, C. & Curtarolo, S. Machine learning for alloys. *Nat. Rev. Mater.* **6**, 1–26 (2021).
31. Heiles, S. & Johnston, R. L. Global optimization of clusters using electronic structure methods. *Int. J. Quantum Chem.* **113**, 2091–2109 (2013).
32. Jäger, M., Schäfer, R. & Johnston, R. L. First principles global optimization of metal clusters and nanoalloys. *Adv. Phys.: X* **3**, 1516514 (2018).
33. Baletto, F. Structural properties of sub-nanometer metallic clusters. *J. Phys.: Condens. Matter* **31**, 113001 (2019).
34. Thorn, A., Rojas-Nunez, J., Hajinazar, S., Baltazar, S. E. & Kolmogorov, A. N. Toward ab initio ground states of gold clusters via neural network modeling. *J. Phys. Chem. C.* **123**, 30088–30098 (2019).
35. Gubaev, K., Podryabinkin, E. V., Hart, G. L. & Shapeev, A. V. Accelerating high-throughput searches for new alloys with active learning of interatomic potentials. *Comput. Mater. Sci.* **156**, 148–156 (2019).
36. Ibarra-Hernández, W. et al. Structural search for stable Mg–Ca alloys accelerated with a neural network interatomic model. *Phys. Chem. Chem. Phys.* **20**, 27545–27557 (2018).
37. Deringer, V. L., Pickard, C. J. & Csányi, G. Data-driven learning of total and local energies in elemental boron. *Phys. Rev. Lett.* **120**, 156001 (2018).
38. Huang, S.-D., Shang, C., Kang, P.-L. & Liu, Z.-P. Atomic structure of boron resolved using machine learning and global sampling. *Chem. Sci.* **9**, 8644–8655 (2018).
39. Podryabinkin, E. V., Tikhonov, E. V., Shapeev, A. V. & Oganov, A. R. Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. *Phys. Rev. B* **99**, 1–7 (2019).
40. Yang, Q. et al. Hard and superconducting cubic boron phase via swarm-intelligence structural prediction driven by a machine-learning potential. *Phys. Rev. B* **103**, 024505 (2021).
41. Deringer, V. L. & Csányi, G. Machine learning based interatomic potential for amorphous carbon. *Phys. Rev. B* **95**, 094203 (2017).
42. Deringer, V. L., Csányi, G. & Proserpio, D. M. Extracting crystal chemistry from amorphous carbon structures. *ChemPhysChem* **18**, 873–877 (2017).
43. Bartók, A. P., Kermode, J., Bernstein, N. & Csányi, G. Machine learning a general-purpose interatomic potential for silicon. *Phys. Rev. X* **8**, 041048 (2018).
44. Deringer, V. L., Proserpio, D. M., Csányi, G. & Pickard, C. J. Data-driven learning and prediction of inorganic crystal structures. *Faraday Discuss.* **211**, 45–59 (2018).
45. Deringer, V. L., Pickard, C. J. & Proserpio, D. M. Hierarchically structured allotropes of phosphorus from data-driven exploration. *Angew. Chem. Int. Ed.* **59**, 15880–15885 (2020).
46. Hajinazar, S., Shao, J. & Kolmogorov, A. N. Stratified construction of neural network based interatomic models for multicomponent materials. *Phys. Rev. B* **95**, 014114 (2017).
47. Sen, R. & Johari, P. Understanding the lithiation of the Sn anode for high-performance Li-ion batteries with exploration of novel Li-Sn compounds at ambient and moderately high pressure. *ACS Appl. Mater. Interfaces* **9**, 40197–40206 (2017).
48. Li, K. et al. From chemistry to mechanics: bulk modulus evolution of Li-Si and Li-Sn alloys via the metallic electronegativity scale. *Phys. Chem. Chem. Phys.* **15**, 17658–17663 (2013).
49. Zhang, P. et al. A first principles study of the mechanical properties of Li–Sn alloys. *RSC Adv.* **5**, 36022–36029 (2015).
50. Mayo, M. & Morris, A. J. Structure prediction of Li-Sn and Li-Sb intermetallics for lithium-ion batteries anodes. *Chem. Mater.* **29**, 5787–5795 (2017).
51. Natesan, K. & Ruther, W. Fabrication and properties of a tin–lithium alloy. *J. Nucl. Mater.* **307**, 743–748 (2002).
52. Li, D., Fürtauer, S., Flandorfer, H. & Cupid, D. Thermodynamic assessment and experimental investigation of the Li–Sn system. *Calphad* **47**, 181–195 (2014).
53. Reichmann, T. L., Li, D. & Cupid, D. M. Heat capacities and an updated thermodynamic model for the Li–Sn system. *Phys. Chem. Chem. Phys.* **20**, 22856–22866 (2018).
54. Kolmogorov, A. et al. New superconducting and semiconducting Fe-B compounds predicted with an ab initio evolutionary search. *Phys. Rev. Lett.* **105**, 217003 (2010).
55. Kolmogorov, A., Shah, S., Margine, E., Kleppe, A. & Jephcoat, A. Pressure-driven evolution of the covalent network in CaB_6. *Phys. Rev. Lett.* **109**, 075501 (2012).
56. Gou, H. et al. Discovery of a superhard iron tetraboride superconductor. *Phys. Rev. Lett.* **111**, 157002 (2013).
57. Fabbris, G. et al. Complex pressure-temperature structural phase diagram of the honeycomb iridate Cu_2IrO_3. *Phys. Rev. B* **104**, 014102 (2021).
58. Sun, F. et al. Electronic and structural response to pressure in the hyperkagome-lattice $Na_3Ir_3O_8$. *Phys. Rev. B* **98**, 085131 (2018).
59. Trimarchi, G., Freeman, A. J. & Zunger, A. Predicting stable stoichiometries of compounds via evolutionary global space-group optimization. *Phys. Rev. B* **80**, 092101 (2009).
60. Zhu, Q., Sharma, V., Oganov, A. R. & Ramprasad, R. Predicting polymeric crystal structures by evolutionary algorithms. *J. Chem. Phys.* **141**, 154102 (2014).
61. Pavone, P., Baroni, S. & de Gironcoli, S. *α-β* phase transition in tin: a theoretical study based on density-functional perturbation theory. *Phys. Rev. B* **57**, 10421 (1998).
62. Legrain, F. & Manzhos, S. Understanding the difference in cohesive energies between alpha and beta tin in DFT calculations. *AIP Adv.* **6**, 045116 (2016).
63. Mehl, M. J. et al. Tin-pest problem as a test of density functionals using high-throughput calculations. *Phys. Rev. Mater.* **5**, 083608 (2021).
64. Wang, L., Maxisch, T. & Ceder, G. Oxidation energies of transition metal oxides within the GGA+U framework. *Phys. Rev. B* **73**, 195107 (2006).
65. de Andrade, A. M., Kullgren, J. & Broqvist, P. Improving the transferability of density functional theory predictions through correlation analysis: Structural and energetic properties of Ni *X* alloys (*X*= C, Si, Ge, and Sn). *Phys. Rev. B* **105**, 085127 (2022).
66. Kolmogorov, A. N. & Crespi, V. H. Registry-dependent interlayer potential for graphitic systems. *Phys. Rev. B* **71**, 235415 (2005).
67. Curtarolo, S., Morgan, D. & Ceder, G. Accuracy of ab initio methods in predicting the crystal structures of metals: a review of 80 binary alloys. *Calphad* **29**, 163–211 (2005).
68. Shao, J., Beaufils, C. & Kolmogorov, A. N. Ab initio engineering of materials with stacked hexagonal tin frameworks. *Sci. Rep.* **6**, 1–8 (2016).
69. Jing, Q.-M. et al. Evidence of polymorphic transformations of Sn under high pressure. *Chin. Phys. B* **25**, 120702 (2016).
70. Cheng, S., Huang, C.-M. & Pecht, M. A review of lead-free solders for electronics applications. *Microelectron. Reliab.* **75**, 77–95 (2017).
71. Salamat, A. et al. High-pressure structural transformations of Sn up to 138 GPa: Angle-dispersive synchrotron x-ray diffraction study. *Phys. Rev. B* **88**, 104104 (2013).
72. Wang, X.-L. et al. Nanospheres of a new intermetallic $FeSn_5$ phase: synthesis, magnetic properties and anode performance in Li-ion batteries. *J. Am. Chem. Soc.* **133**, 11213–11219 (2011).
73. Pettifor, D. G. et al. Bonding and Structure of Molecules and Solids (Oxford university press, 1995).
74. Guillaume, C. L. et al. Cold melting and solid structures of dense lithium. *Nat. Phys.* **7**, 211–214 (2011).
75. Hutcheon, M. & Needs, R. Structural and vibrational properties of lithium under ambient conditions within density functional theory. *Phys. Rev. B* **99**, 014111 (2019).
76. Zhang, W. et al. Unexpected stable stoichiometries of sodium chlorides. *Science* **342**, 1502–1505 (2013).
77. Zhang, W. et al. Stability of numerous novel potassium chlorides at high pressure. *Sci. Rep.* **6**, 1–6 (2016).
78. Sonia, F. J. et al. Understanding the Li-storage in few layers graphene with respect to bulk graphite: experimental, analytical and computational study. *J. Mater. Chem. A* **5**, 8662–8679 (2017).
79. Goward, G., Taylor, N., Souza, D. & Nazar, L. The true crystal structure of $Li_{17}M_4$ (M= Ge, Sn, Pb)–revised from $Li_{22}M_5$. *J. Alloy. Compd.* **329**, 82–91 (2001).
80. Genser, O. & Hafner, J. Structure and bonding in crystalline and molten Li-Sn alloys: a first-principles density-functional study. *Phys. Rev. B* **63**, 144204 (2001).
81. Gasior, W., Moser, Z. & Zakulski, W. Thermodynamic studies and the phase diagram of the Li-Sn system. *J. Non-Crystal. solids* **205**, 379–382 (1996).
82. Blum, V. & Zunger, A. Mixed-basis cluster expansion for thermodynamics of bcc alloys. *Phys. Rev. B* **70**, 155108 (2004).
83. Kadkhodaei, S. & Muñoz, J. A. Cluster expansion of alloy theory: a review of historical development and modern innovations. *JOM* **73**, 3326–3346 (2021).
84. Müller, W. Darstellung und struktur der phase Li_7Sn_3/Preparation und crystal structure of Li_7Sn_3. *Z. f.ür. Naturforsch. B* **29**, 304–311 (1974).
85. Müller, W. & Schäfer, H. Die kristallstruktur der phase LiSn: the crystal structure of LiSn. *Z. f.ür. Naturforsch. B* **28**, 246–248 (1973).
86. Blase, W. & Cordier, G. Crystal structure of *β*-lithium stannide, *β*-LiSn. *Z. f.ür. Kristallographie-Crystalline Mater.* **193**, 317–318 (1990).
87. Hansen, D. A. & Chang, L. J. Crystal structure of Li_2Sn_5. *Acta Crystallogr. Sect. B* **25**, 2392–2395 (1969).

88. Sorella, S., Casula, M., Spanu, L. & Dal Corso, A. Ab initio calculations for the β-tin diamond transition in silicon: comparing theories with experiments. *Phys. Rev. B* **83**, 075119 (2011).
89. Ravelo, R. & Baskes, M. Equilibrium and thermodynamic properties of grey, white, and liquid tin. *Phys. Rev. Lett.* **79**, 2482–2485 (1997).
90. Gaál-Nagy, K. et al. Temperature and dynamical effects on the high-pressure cubic-diamond to beta-tin phase transition in Si and Ge. *Phys. Status Solidi (b)* **211**, 275–280 (1999).
91. Christensen, N. E. & Methfessel, M. Density-functional calculations of the structural properties of tin under pressure. *Phys. Rev. B* **48**, 5797–5807 (1993).
92. Kubota, Y., Escaño, M. C. S., Nakanishi, H. & Kasai, H. Crystal and electronic structure of $Li_{15}Si_4$. *J. Appl. Phys.* **102**, 053704 (2007).
93. Johnson, Q., Smith, G. S. & Wood, D. The crystal structure of $Li_{15}Ge_4$. *Acta Crystallogr.* **18**, 131–132 (1965).
94. Zeng, Z. et al. A novel phase of $Li_{15}Si_4$ synthesized under pressure. *Adv. Energy Mater.* **5**, 1500214 (2015).
95. Stournara, M. E., Guduru, P. R. & Shenoy, V. B. Elastic behavior of crystalline Li–Sn phases with increasing Li concentration. *J. Power Sources* **208**, 165–169 (2012).
96. Kolmogorov, A., Hajinazar, S., Angyal, C., Kuznetsov, V. & Jephcoat, A. Synthesis of a predicted layered LiB via cold compression. *Phys. Rev. B* **92**, 144110 (2015).
97. Sandoval, E. D., Hajinazar, S. & Kolmogorov, A. N. Stability of two-dimensional BN-Si structures. *Phys. Rev. B* **94**, 094105 (2016).
98. Xu, Y. et al. Large-gap quantum spin hall insulators in tin films. *Phys. Rev. Lett.* **111**, 136804 (2013).
99. Wang, J., Raistrick, I. & Huggins, R. A. Behavior of some binary lithium alloys as negative electrodes in organic solvent-based electrolytes. *J. Electrochem. Soc.* **133**, 457 (1986).
100. Courtney, I. A., Tse, J. S., Mao, O., Hafner, J. & Dahn, J. R. Ab initio calculation of the lithium-tin voltage profile. *Phys. Rev. B* **58**, 15583–15588 (1998).
101. Tran, T. T. & Obrovac, M. N. Alloy negative electrodes for high energy density metal-ion cells. *J. Electrochem. Soc.* **158**, A1411 (2011).
102. Stratford, J. M. et al. Investigating sodium storage mechanisms in tin anodes: a combined pair distribution function analysis, density functional theory, and solid-state NMR approach. *J. Am. Chem. Soc.* **139**, 7273–7286 (2017).
103. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
104. Kresse, G. & Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251–14269 (1994).
105. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15 – 50 (1996).
106. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
107. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
108. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
109. Langreth, D. C. & Mehl, M. J. Beyond the local-density approximation in calculations of ground-state electronic properties. *Phys. Rev. B* **28**, 1809–1834 (1983).
110. Ceperley, D. M. & Alder, B. J. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.* **45**, 566 (1980).
111. Perdew, J. P. & Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **23**, 5048 (1981).
112. Sun, J., Ruzsinszky, A. & Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **115**, 036402 (2015).
113. Monkhorst, H. J. & Pack, J. D. Special points for brillouin-zone integrations. *Phys. Rev. B* **13**, 5188–5192 (1976).
114. Togo, A. & Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **108**, 1–5 (2015).
115. Towns, J. et al. XSEDE: Accelerating scientific discovery. *Comput. Sci. Eng.* **16**, 62–74 (2014).
116. Belsky, A., Hellenbrandt, M., Karen, V. L. & Luksch, P. New developments in the inorganic crystal structure database (ICSD): Accessibility in support of materials research and design. *Acta Crystallogr. Sect. B: Struct. Sci.* **58**, 364–369 (2002).
117. Momma, K. & Izumi, F. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *J. Appl. Crystallogr.* **44**, 1272–1276 (2011).

ACKNOWLEDGEMENTS

We acknowledge the NSF support (Award No. DMR-1821815) and the Extreme Science and Engineering Discovery Environment computational resources[115] (NSF Award No. ACI-1548562, Project No. TG- PHY190024).

AUTHOR CONTRIBUTIONS

A.N.K. conceptualized the work. All authors contributed to running the simulations, analyzing the results, and writing the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-022-00825-4.

Correspondence and requests for materials should be addressed to Aleksey N. Kolmogorov.

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.