J Mater Sci

Computational Materials Design

Review of computational approaches to predict the thermodynamic stability of inorganic solids

Christopher J. Bartel[1,*]

[1] *Department of Materials Science and Engineering, University of California, Berkeley, USA*

Received: 28 September 2021
Accepted: 12 January 2022

ABSTRACT

Improvements in the efficiency and availability of quantum chemistry codes, supercomputing centers, and open materials databases have transformed the accessibility of computational materials design approaches. Thermodynamic stability predictions play a central role in the efficacy of these approaches and should be considered carefully. This review covers the fundamentals of calculating thermodynamic stability using first-principles methods. Stability is delineated into two main topics—stability with respect to decomposition into competing phases and stability with respect to phase transition into alternative structures at fixed composition. For each topic, a summary of the state-of-the-art is provided along with a tutorial overview of practical considerations. The application of machine learning to both kinds of stability predictions is also covered. Finally, the limitations of thermodynamic stability predictions are discussed within the context of predicting the synthesizability of materials.

Introduction

The virtual design and discovery of new inorganic crystalline solids using first-principles calculations is a core discipline in materials science research [1–3]. Quantum chemical approaches such as density functional theory (DFT) [4] are commonly used to calculate the properties of known and hypothetical materials, shedding light on their underlying composition-structure–property relationships and advancing applications including batteries [5], photovoltaics [6], thermoelectrics [7], catalysts [8], quantum materials [9], and many others. Open materials databases [10–18], which together house DFT-calculated properties for millions of inorganic crystals, significantly aid this process by allowing any researcher to leverage the computational expense of others for projects of their interest. Along with these massive open data sets, machine learning (ML) has emerged as a complementary tool to further alleviate computational expense or assist in the exploration of novel materials [19–21]. The success of high-throughput materials discovery [22] and the inverse design of novel materials with targeted properties [23] depends critically on the calculation of each

Handling Editor: M. Grant Norton.

Address correspondence to E-mail: cbartel@berkeley.edu

https://doi.org/10.1007/s10853-022-06915-4

Published online: 01 February 2022

candidate material's thermodynamic stability [24]. This review will provide a tutorial introduction into the role of thermodynamic stability in materials design, an overview of existing approaches to predict stability using DFT and ML, and some concluding thoughts on moving beyond the prediction of stability and towards the prediction of synthesizability.

Defining thermodynamic stability

An inorganic crystalline solid material is defined by two factors — its composition (what elements are in it and in what proportions) and its structure (how the atoms of those elements arrange themselves into a periodic crystal). For most of this review, we will consider ground-state DFT calculations, which do not include temperature effects (e.g., entropic effects, thermal expansion), and accordingly the thermodynamic potential is simply the total (internal) energy, E. A material is thermodynamically stable under a given set of conditions if its energy cannot be lowered by rearranging its atoms. Energy lowering can arise from two distinct mechanisms: 1) phase separation (decomposition) into competing materials with the same average composition or 2) phase transition to an alternative crystal structure (polymorph) at fixed composition. These concepts will be introduced in this section and discussed in more detail with relevant approaches and examples in the subsequent sections.

All possible phase separation and phase transition reactions must be thermodynamically unfavorable for a material to be deemed stable at a specified set of conditions. The choice of (a fixed set of) conditions (temperature, pressure, open/closed system, chemical potentials, etc.) that describe a system determine the appropriate thermodynamic potential with which to evaluate the energy, and material stability can be dynamic with respect to this potential [25–29]. As a simple example, high-symmetry crystal structures (e.g., cubic) are often stabilized at elevated temperatures relative to alternative polymorphs. Alternatively, a metal oxide with a strongly oxidized transition metal may be stable at low temperature but at high temperature, becomes unstable with respect to phase separation into O_2 and an alternative metal oxide that places the transition metal in a more reduced state.

Once a set of conditions are specified, and the energies of all relevant materials are computed with the appropriate thermodynamic potential, stability can be assessed with respect to phase separation and phase transition. Stability with respect to phase transition is determined by comparing the energy of the material of interest to the energies of all crystalline polymorphs having the same composition. The crystal structure having the lowest energy is often called the ground-state polymorph and is stable with respect to phase transition. The magnitude of polymorphic (in)stability will be referred to in this work as ΔE_{gs}, where $\Delta E_{gs} = 0$ for the ground-state structure and $\Delta E_{gs} > 0$ for non-ground-state polymorphs. For example, consider a hypothetical composition, A_2X_5 (where A and X are arbitrary chemical elements) that has two possible crystal structures — α and β, where α is the ground state and β is a metastable polymorph. The polymorphic instability of the β phase can be obtained by simply computing the difference in total energies of the two phases: $\Delta E_{gs,\beta} = E_\beta - E_\alpha$.

Stability with respect to phase separation is then assessed for all ground-state polymorphs across the relevant chemical space of interest using the convex hull formalism (Fig. 1). Similar in spirit to Gibbs energy minimization [30, 31], the convex hull analysis assesses whether a given material can lower its energy by decomposing into a linear combination of materials having the same average composition as the material of interest. Consider again a candidate material with the composition A_2X_5. The formation energy, ΔE_f, of A_2X_5 is first obtained by referencing to the elemental energies:

$$\Delta E_{f,A_2X_5} = E_{A_2X_5} - 2E_A - 5E_X$$

In other words, ΔE_f is the reaction energy for forming a compound from its elements (e.g., $2A + 5X \rightarrow A_2X_5$), and $\Delta E_f > 0$ indicates the compound is unstable with respect to its elemental phases. This quantity is sometimes also referred to as the "mixing energy", or the extent to which the energy is lowered ($\Delta E_f \leq 0$) or increased ($\Delta E_f > 0$) by mixing stoichiometric amounts of components A and B. It should be noted that $\Delta E_f \leq 0$ does not indicate that a material is thermodynamically stable, only that it will not phase separate into its elemental components. This topic will be discussed in more detail in ''From formation to decomposition reactions''.

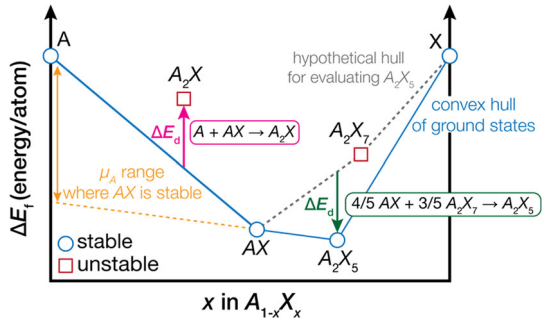

Figure 1 Illustrating stability within a convex hull phase diagram. Note that for clarity, only ground-state phases at each composition are shown here. There could be many hypothetical non-ground-state polymorphs at any given composition. The solid blue line is the convex hull. Points lying on the hull (blue circles) indicate thermodynamically stable phases. Points lying above the hull are thermodynamically unstable with respect to phase separation. The boxed reactions indicate the reactions that define ΔE_d for A_2X and A_2X_5. Note that these reactions must be normalized per atom (e.g., by converting AX to $A_{0.5}X_{0.5}$) to obtain ΔE_d. The dashed gray line refers to the hypothetical convex hull used to determine the decomposition reaction and energy for the stable phase, A_2X_5. The dashed orange line is the tangent to the convex hull at composition AX and provides a lower bound on the chemical potential of A, μ_A, where AX would still be on the convex hull (stable).

Once ΔE_f is similarly obtained for all ground-state polymorphs in the A–X chemical space (e.g., A_2X, AX, A_2X_7, etc.), these formation energies are produced as a function of some extensive property (or properties) of the system, usually the molar composition of $N-1$ elements in a chemical space having N unique elements —e.g., x in $A_{1-x}X_x$. The lower convex envelope that joins these points in ΔE_f–x space is called the convex hull of stability and is shown by the solid blue line in Fig. 1. Because the convex hull is defined over fractional composition space (e.g., $x = 2/3$ for AX_2), each compound should be represented with a basis of 1 mol of atoms per compound, and ΔE_f must therefore be normalized per atom in each formula. In mathematics, a convex hull is the object that bounds the extreme points in a set (roughly speaking). If some point, Z, lies on an N-dimensional convex hull, where $N-1$ dimensions are independent (e.g., chemical composition) and the Nth dimension is dependent (e.g., formation energy), then no linear combination of points constrained to have the same average value as Z for all $N-1$ independent dimensions can possibly have a more extreme value than Z in the Nth dimension. The convex hull is a convenient tool for thermodynamics because materials that lie on the convex hull are thermodynamically stable with respect to phase separation, whereas materials that lie above the hull are unstable because there is a linear combination of alternative compositions through which their energy can be lowered. In thermodynamics, we care only about the "lower" convex hull as points that lie above the $\Delta E_f = 0$ (or equivalent reference state) are necessarily unstable.

The magnitude of (in)stability with respect to phase separation is called the decomposition energy, ΔE_d. For unstable materials that lie above the convex hull, $\Delta E_d > 0$ and is equivalent to the vertical distance in formation energy space between that material and the hull, as shown for A_2X in Fig. 1. ΔE_d is the reaction energy for forming A_2X from the linear combination of neighbors in composition space (i.e., the energy for the hypothetical decomposition reaction: $A + AX \rightarrow A_2X$), typically normalized per atom in the compound of interest (e.g., divided by the three atoms that comprise A_2X). For stable materials that lie on the convex hull, $\Delta E_d \leq 0$ and is equivalent to the vertical distance in formation energy space between that material and a hypothetical convex hull drawn without the material of interest (shown as the dashed gray line in Fig. 1 for evaluating A_2X_5). Note that unstable materials can appear in the decomposition reaction for stable materials because we are considering a hypothetical convex hull that excludes the stable material of interest. In the Fig. 1 example, the unstable material, A_2X_7, appears in the decomposition reaction for the stable material, A_2X_5, because A_2X_7 would be on the hull if A_2X_5 did not exist.

$|\Delta E_d|$ therefore provides the positive change in ΔE_f that a stable material can tolerate and remain stable or the negative change in ΔE_f required for an unstable material to become stable (all other formation energies being fixed). The related energy above the convex hull, ΔE_{hull}, is often reported as the metric for thermodynamic stability with respect to phase separation or transition [10]. ΔE_{hull} and ΔE_d are equivalent for unstable materials, but for stable materials, $\Delta E_{hull} = 0$, whereas ΔE_d can be any value ≤ 0. In contrast to ΔE_{hull} which only indicates that a stable material is stable, ΔE_d quantifies how stable that material is, providing useful information with respect to uncertainty in the assessment of stability as well as plausibility of synthesis. Although illustrated here for a binary chemical space, A–X, this approach generalizes exactly to chemical spaces

comprised of any number of elements. Practical considerations for the calculation of ΔE_f, ΔE_d, and ΔE_{gs} will be discussed in detail in the subsequent sections.

Stability with respect to phase separation

In a typical materials discovery, design, or screening effort, a large set of candidate materials are initially considered. Certain filters are then applied to systematically focus the set of candidate materials to those most attractive for a given application. Regardless of application, an essential filter is one based on stability with respect to phase separation (ΔE_d). While it's clear that some ΔE_d filter is needed to avoid including many unstable (and therefore unlikely-to-be-synthesized) materials as candidates, it is not so clear what the ΔE_d cutoff should be to retain materials as potentially interesting candidates. In an ideal (but less interesting) world where 1) our calculations were exact and 2) metastable materials could not be made, we would only consider materials as viable candidates if $\Delta E_d \leq 0$. However, we know that our DFT calculations are approximations of reality and many thousands of materials that are calculated to be unstable with DFT at 0 K have been synthesized [32]. This arises from both uncertainties in the DFT-calculated thermodynamics [33] and the (lack of) conditions considered in these calculations (temperature, pressure, etc.). Still, ΔE_d provides essential information regarding the thermodynamic driving forces to form (or decompose) desired materials and is an essential tool in the high-throughput prediction of novel materials. Beyond using ΔE_d to identify synthesizable materials, the magnitude of ΔE_d is also indicative of the relative ground-state electronic interactions in materials. Because ΔE_d arises from comparing ΔE_f (which arises from internal energies, E), energy differences arise only from chemical bonding effects, and not from, for example, entropic effects.

Formation energy predictions

For thermodynamic calculations of inorganic crystals, any density functional that yields total energies can be used, but because the generalized gradient approximation (GGA) is a standard approximation to DFT [34], some practical points will focus on this approach. ΔE_f is the starting point for determining ΔE_d and accurate predictions of ΔE_f require that the energy differences between compounds and their elements be accurately captured. Much of the success of local and semi-local density functionals in capturing materials thermodynamics arises because of convenient error cancellation in total energies. That is, there may be intrinsic errors in calculated total energies, but if these errors are systematic, the calculation of relative energies will have much smaller errors. However, the extent of error cancellation is related to the extent that the materials being compared are similar. For example, it was recognized many years ago that GGA overbinds the O_2 molecule, resulting in μ_O that is too negative (where $\mu_O = E_O$ and is the chemical potential of oxygen gas) [35]. ΔE_f for solid oxides requires comparing the energies of the solid to O_2 (and any other elements in the material). The extent of overbinding for molecular O_2 is dissimilar to any overbinding in cation-oxygen bonds occurring in a crystal, so GGA exhibits large errors for ΔE_f of oxides. Fortunately, this error can be easily corrected by modifying the elemental reference energy for oxygen, μ_O. Wang et al. fit a simple correction to μ_O for a set of six binary main group metal oxides (Li_2O, Na_2O, MgO, CaO, Al_2O_3, SiO_2) by comparing GGA-calculated ΔE_f to experimentally measured enthalpies of formation, ΔH_f, at standard conditions (298 K, 1 atm) [36].

This concept has also been generalized such that the reference energies of all elements could be fit to maximize consistency with experimental measurements [37, 38]. In this approach, a set of compounds with experimentally known ΔH_f (from tabulated calorimetry data [39–42]) are calculated with a given DFT approximation (e.g., GGA). The reference energies, μ_i, for the elements of all compounds are then simultaneously optimized to minimize the sum of squared differences between calculated ΔE_f and measured ΔH_f. ΔE_f for an arbitrary compound can be written as follows, where n_i is the number of moles of element i in the compound, and $\delta\mu_i$ is the correction to the total energy of element i that is to be fit.

$$\Delta E_{f,A_aX_x\ldots} = E_{A_aX_x\ldots} - \sum_{i}^{A,X,\ldots} n_i(\mu_i + \delta\mu_i) \quad (1)$$

Given a set of compounds with calculated ΔE_f and measured ΔH_f, we get an overdetermined system of linear equations of the form of Eq. 1, that can be

solved to minimize the disagreement between ΔE_f and ΔH_f by modification of each $\delta\mu_i$. Stevanović et al. showed that the error in calculated ΔE_f (relative to measured ΔH_f) can be reduced from 250 meV/atom with pure GGA to 54 meV/atom with elemental reference energies fit in this fashion for 252 binary solids [38]. This approach was used by Kirklin et al. for a larger set of 1,670 compounds, with errors decreasing from 136 meV/atom with no fitted reference energies to 81 meV/atom after fitting all elemental reference energies [12]. For context, "chemical accuracy" is sometimes quoted as resolution within 1 kcal/mol ($\sim$ 22 meV/atom for an equimolar binary compound). However, this is simply a rule-of-thumb meant to estimate the expected uncertainty of experimental measurements of ΔH_f, and as-such, should not be considered too quantitatively. An "acceptable" resolution strongly depends upon the desired outcome of the calculations.

Beyond accounting for systematic errors with the treatment of elemental phases, ΔE_f prediction accuracy is also dependent upon the choice of functional. As two examples, including a Hubbard U correction (GGA + U) can improve the description of systems with transition metals [43] and more sophisticated functionals, such as the meta-GGA SCAN functional [44], can generally improve calculated ΔE_f and mitigate the need for fitted reference energies or the U correction by more accurately determining the energy of diversely bonded systems (e.g., by correcting O_2 overbinding) [45, 46]. While there are many DFT approaches available for calculating total and formation energies, a consistent approach must be used for all competing phases in a given chemical space in order to yield meaningful stability predictions (ΔE_d). In this respect, it is often advantageous to make use of the abundant set of materials available in open databases that already provide calculated ΔE_f. The stability of a new candidate material can then be obtained by calculating only that material and comparing to the materials in a given database as long as the procedure used to calculate the new candidate and that used by the database are the same (or proven to be compatible). Instead, if an alternative approach (functional, U correction, etc.) is used for a new candidate material, then all relevant competing phases should be recalculated with this same alternative approach in order to consistently determine ΔE_d for this candidate. This generally precludes the use of very high level (e.g., beyond-DFT [47]) approaches to calculate ΔE_d for anything but the simplest chemical spaces. In principle, one could also develop a scheme to make disparate approaches compatible with one another, as is done in the scheme used by Materials Project for mixing GGA and GGA + U calculations [43].

From formation to decomposition reactions

Because DFT-calculated total internal energies are not meaningful on an absolute scale, ΔE_f is the principal thermodynamic input to the convex hull construction, and benchmarking efforts traditionally focus on this quantity. However, it is important for the purposes of assessing stability that we understand how our chosen calculation procedure affects ΔE_d, which is the quantity governing material stability. A comparison of ΔE_d and ΔE_f for $\sim$ 85,000 compounds in Materials Project shows clearly that these two quantities have a highly nonlinear relationship (Fig. 2), so we should not necessarily expect that benchmarking on ΔE_f is transferable to ΔE_d. Further, though it is sometimes stated in the literature that because $\Delta E_f < 0$, a material might be stable, available data does not support this. Of the 85,014 ground-state materials ($\Delta E_{gs} = 0$) drawn from Materials Project and shown

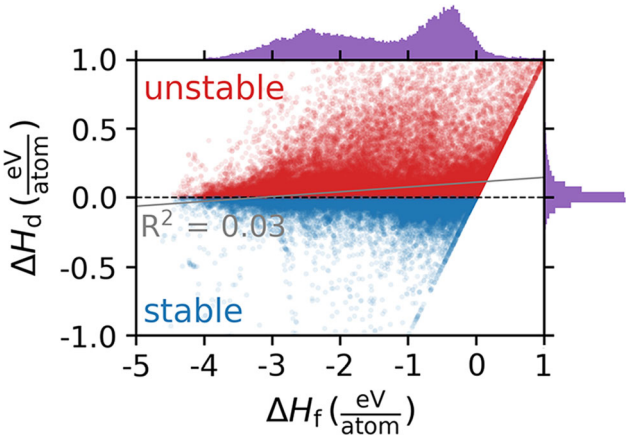

Figure 2 The decomposition enthalpy, ΔH_d, shown against the formation enthalpy, ΔH_f, (equivalently ΔE_d and ΔE_f) for 85,014 ground-state entries in Materials Project, indicating little linear correlation between the two quantities. The strong linear correlation that is present at $\Delta H_d = \Delta H_f$ arises for chemical spaces that contain only one compound ($\sim$ 3% of compounds). These compounds were excluded from the correlation coefficient, R^2, determination. A normalized histogram of ΔH_f (ΔH_d) is shown above (along the right side of). Reproduced with permission from Ref. [48].

in Fig. 2, 46,582 (~ 55%) have $\Delta E_f < 0$ but are unstable with respect to phase separation ($\Delta E_d > 0$), emphasizing the importance of computing stability relative to competing compounds (and not just elements).

In the same way that there is error cancellation when comparing total energies, E, in the determination of ΔE_f, there is also error cancellation when comparing formation energies in the determination of ΔE_d. This was shown in the work of Hautier et al. [49], where they studied the formation of ternary oxides from their binary oxide counterparts, comparing GGA and experimental reaction energies for reactions of the form:

$$A_aO_x + M_mO_y \rightarrow A_aM_mO_{x+y}$$

While these are not necessarily the actual decomposition reactions for each ternary oxide, the 135 evaluated reactions provide a reasonable set over which to understand the extent of error cancellation that can occur for decomposition-like reactions that do not contain elemental phases. For these reactions, all elemental reference energies cancel such that the associated reaction energies can be computed equivalently with either formation energies or total energies for each compound. The mean absolute error (MAE) for GGA compared to experimental reaction enthalpies for these reactions was found to be only 24 meV/atom, an order of magnitude lower than the GGA error on ΔE_f reported in Ref. [38]. This result encourages the notion that DFT errors mostly cancel one another for chemically similar materials, which is of critical importance for stability predictions. However, it should be noted that the reactions studied by Hautier et al. are not necessarily the decomposition reactions for each ternary oxide as a ternary compound can compete for stability with a mixture of elemental, binary, and ternary phases, not just the isovalent binary oxides evaluated in this study.

A more systematic evaluation of decomposition reactions was performed by Bartel et al., where experimental decomposition enthalpies, ΔH_d, were determined for 1,012 compounds having tabulated ΔH_f [46]. These experimentally obtained ΔH_d were then compared to DFT-calculated ΔE_d using GGA and SCAN. For the purposes of assessing the nature of errors (and error cancellation) for ΔE_d, decomposition reactions were divided into three types: (i) Type 1 reactions where a compound competes only with elemental phases and $\Delta E_f = \Delta E_d$; (ii) Type 2 reactions, similar to those studied by Hautier et al., where compounds compete only with other compounds and elemental reference energies cancel; and (iii) Type 3 reactions where both compounds and elements appear in the stability-defining decomposition reaction. Type 1 reactions will only occur for materials that are the sole compound in their chemical space with $\Delta E_f < 0$ and are therefore quite rare. Type 3 reactions occur for compounds that are near the elemental boundaries of a phase diagram (i.e., at high or low values of x in Fig. 1). These decomposition reactions are interesting from a synthetic perspective because the value of ΔE_d can be modulated by controlling the chemical potential of the element that participates in the decomposition reaction [29]. Type 2 reactions are the most common, accounting for 63% of the ~ 50,000 decomposition reactions involving compounds tabulated in Materials Project. The prevalence of Type 2 reactions increases as the number of elements in a material (and therefore the number of competing compounds in a chemical space) increases. For 231 Type 2 decomposition reactions, GGA and SCAN were found to have nearly identical agreement with experiment, exhibiting MAE of 34 meV/atom for ΔE_d calculated using either functional. For 415 Type 3 reactions, calculated errors were higher because of the presence of elemental phases, and SCAN improved slightly upon GGA (MAE of 73 meV/atom compared to 90 meV/atom). Notably, fitting elemental reference energies to ΔE_f did not substantially improve ΔE_d predictions by either functional, though fitted reference energies may still prove useful for certain element-forming/consuming reactions such as those considered for defect formation [50].

A practical consideration in the determination of ΔE_d is which competing phases to include in the analysis. An exhaustive approach would consider all known competing phases in each chemical space of interest (e.g., by querying the Inorganic Crystal Structure Database (ICSD) [51]). This approach is generally tractable if the desired calculation procedure for any new material is compatible with the approach used by one of the open materials databases, for which most of the ICSD has already been calculated. However, if an alternative procedure (e.g., functional) is needed to accurately describe the materials of interest, it may be impractical to recompute all known competing phases within this

compatible approach. For example, if one is pursuing new materials in the Sr-Fe-Mo–O chemical space, the ICSD contains > 100 unique materials that, in principle, should be considered for determining the stability of the new candidate materials. If the computational search spans many ternary, quaternary, or higher order spaces, the number of competing phases that need to be calculated grows tremendously. An attractive simplification of this problem would be to only consider the common binary oxides, which often comprise the precursors used to synthesize higher order materials. However, based on Materials Project data, ternary (or higher order) compounds are present in the decomposition reaction for 42% of ternary compounds and 91% of quaternary compounds [46]. The approximation of ΔE_d as the reaction energy from the binaries would therefore represent a lower bound on ΔE_d as the introduction of additional competing phases can only increase ΔE_d. In situations where it is impractical to apply a chosen calculation procedure to all possible competing phases, materials should be prioritized based on ΔE_d (if available at some level of theory) and proximity in composition space to candidate materials of interest.

Alternative thermodynamic potentials

Although most applications of DFT for material stability consider isolated systems in vacuum at 0 K, the effect of temperature and other intensive properties can be captured in a thermodynamic stability analysis. As a simple example, the formation energy at 0 K, ΔE_f, can be mapped into a formation enthalpy, ΔH_f, at some temperature, T, by including the zero-point energy correction to the DFT-calculated total energy and integrating the constant volume specific heat from 0 K to T. Both of these quantities can be obtained from the DFT-calculated phonon density of states [52]. In this case, it has been shown that formation energies (and by consequence, decomposition energies) are generally unaffected by this mapping from 0 K energies to 298 K enthalpies, again due to cancellation of errors [46].

However, if we consider higher temperatures (such as those often required for the synthesis of solids), the role of entropy, S, in governing equilibrium grows linearly as the appropriate thermodynamic potential is the Gibbs energy, $G = H - TS$. Gibbs formation energies, ΔG_f, can differ quite substantially from ΔE_f, especially at high temperature. As an example, for Al_2O_3, $\Delta E_f \approx -3.4$ eV/atom whereas $\Delta G_f(1000\text{ K}) \approx -2.8$ eV/atom [42]. The primary contribution to the temperature dependence of ΔG_f is vibrational entropy [53], which must be considered for both the compound of interest and the elemental reference states. Once again, error cancellation plays a role in the accuracy of ΔG_f and the resulting decomposition energy, ΔG_d. If all species in a given chemical reaction respond to temperature similarly, then ΔE_f (ΔE_d) should be a suitable approximation for ΔG_f (ΔG_d). Within the context of the types of decomposition reactions discussed previously, the energies of Type 1 and 3 reactions will likely be subject to substantial temperature effects whereas Type 2 reactions will not. For example, when only solids are present in the reaction, ΔE_f (ΔE_d) approximates ΔG_f (ΔG_d) [26]. When gases are involved in a given chemical reaction (e.g., in the formation reaction for oxides, nitrides, halides, etc.), it is often assumed that only the temperature-dependence of the gaseous species needs to be considered [25]. However, it was shown in Ref. [26] that an incomplete cancellation of vibrational entropy between the compound (e.g., TiO_2) and the elemental solid (e.g., Ti) leads to significant error in this approximation that grows as a function of temperature. A simple model for the vibrational entropy of solids was proposed in Ref. [26] to readily extend DFT-calculated formation energies to finite temperature, though the accuracy may be more suitable for high-throughput evaluations of T-dependent thermodynamics, rather than detailed examinations of T-dependent phase transitions.

For systems exhibiting disorder (e.g., materials with mixed occupancies), configurational entropy also contributes significantly to ΔG_f. This is the prevailing contribution stabilizing disordered materials (e.g., alloys [54], solid solutions [55], and "high-entropy" compounds [56]) relative to their ordered counterparts. The simplest approach to quantify the configurational entropy in a system is to assume the mixed species behave as an ideal solution. Considering a binary alloy of A and X— $A_{1-x}X_x$ —where A and X exist on the same site in the lattice in equal average occupation, the ideal mixing entropy, S_{mix} is calculated as:

$$S_{mix} = -k_B[xlnx + (1-x)\ln(1-x)]$$

where k_B is Boltzmann's constant. This provides an upper bound on the configurational entropy of $\sim 0.69\ k_B$/atom (~ 0.06 meV/atom/K) for an equimolar solid solution. For comparison, vibrational entropy differences between intermetallic phases are typically ~ 0.1–$0.2\ k_B$/atom [57]. A more sophisticated approach for estimating configurational entropy that accounts for short-range ordering involves cluster expansion and Monte Carlo simulations, which have been reviewed elsewhere [58]. In chemical spaces having partially or fully disordered materials, configurational entropy can play a significant role in defining the convex hull and decomposition energies of all phases in that space and becomes more significant as more species contribute to disorder [59, 60].

In addition to computing the enthalpic and entropic contributions to the Gibbs energy, an alternative method is through the CALculation of PHAse Diagrams (CALPHAD) approach [61]. This approach is particularly applicable to the construction of phase diagrams for multicomponent materials modeled as solid solutions using a compound energy formalism [62]. Using computed or experimentally obtained thermodynamic properties of materials as inputs, multicomponent phase diagrams are obtained by fitting models for the Gibbs energy of constituent (mixed) phases and minimizing the collective Gibbs energy of the system subject to constraints imposed by the laws of thermodynamics. Materials systems that have been extensively studied (e.g., binary intermetallics) are more amenable to this approach because the input data required to fit the relevant Gibbs energy models for these systems are already measured or calculated. For a more comprehensive discussion of the details, applications, and limitations of CALPHAD, the interested reader is referred to the recent overview by Liu [63].

The Gibbs energy is the natural thermodynamic potential for a closed system at fixed temperature and pressure, but materials scientists are often interested in the stability of materials in systems open to exchange with the environment (e.g., O_2 in air during synthesis) [25, 64]. For a system open to some element, X, we can construct a grand potential, Φ:

$$\Phi = G - n_X \mu_X$$

where μ_X is the chemical potential of the reservoir of X, and n_X the number of moles of X in the compound being evaluated. An elemental reservoir may be appropriate to consider for systems in a reactive environment (e.g., air, water, etc.) or interfaced with a reactive species (e.g., a Li metal anode). The grand potential (sometimes referred to as the grand canonical or Landau potential) can be more generally written as:

$$\Phi = G - \sum_i X_i Y_i$$

where X_i is an extensive property (e.g., number of moles of some species) and Y_i its intensive conjugate (e.g., chemical potential of that same species). These grand potentials can be customized to account for the handles that are most relevant to a given system — e.g., partial pressure effects on oxide surfaces [65, 66] or particle size effects during nucleation [27, 67].

Just as with internal energies and Gibbs energies, we can also obtain grand potential formation energies, $\Delta\Phi_f$, by referencing to elemental phases and grand potential decomposition energies, $\Delta\Phi_d$, by performing the convex hull analysis within this potential space. As one example, this approach has been applied extensively in the computational search for solid-state electrolytes for Li-ion batteries (Fig. 3) [68–70]. In these systems, we are often interested in the electrochemical stability window of a candidate solid-state Li^+ conductor when interfaced with a Li reservoir (the electrodes). Performing a convex hull analysis using a grand potential with respect to Li allows the determination of the range of applied voltages (equivalently Li chemical potentials, μ_{Li}, relative to a reference chemical potential, μ^0_{Li}) for which a given material is thermodynamically stable (on the convex hull). This approach is

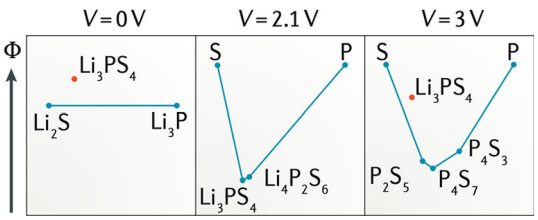

Figure 3 A convex hull construction in grand potential space for a system open to exchange with Li reveals the electrochemical stability window for Li_3PS_4. The grand potential, Φ, is shown along the y-axis (normalized per non-Li atom in Li_3PS_4), the composition of P (relative to S) along the x-axis, with increasing applied voltage (Li chemical potential) moving from the left panel to the right one. Reproduced with permission from Ref. [70].

illustrated for Li_3PS_4 in Fig. 3, where this material is calculated to be electrochemically stable at 2.1 V ($\mu_{Li} = \mu^0_{Li} - 2.1$ eV), unstable with respect to decomposition into $Li_2S + Li_3P$ (− Li) at 0 V, and unstable with respect to $S + P_2S_5$ (+ Li) at 3 V. In the same way that ΔE_d quantifies the driving force for (in)stability in a closed system at 0 K, the magnitude of $\Delta\Phi_d$ provides the driving force for forming certain phases at a specified set of conditions relevant to the synthesis of a material or the operation of a device.

An example of the grand potential approach for more complex thermodynamic systems is the Pourbaix potential, which is applicable for ions in aqueous electrochemical systems. The grand potential for an ion in an aqueous system at a redox potential E^0 takes the form:

$$\Phi = G - n_H\left(\mu_{H^+} - E^0\right) - E^0 Q - n_O\mu_O$$

where Q is the charge of the ion and μ_H and μ_O are constrained by the equilibrium of water and oxygen [27]. The neutral or ionic phase that minimizes Φ at a specified redox potential and pH is the stable phase and therefore the one shown on the Pourbaix diagram [71] at those conditions. Care must be taken in this approach to properly reference energies such that the free energies of solvated ions in solution do not need to be calculated [72].

Stability in terms of chemical potentials

The Pourbaix phase diagram falls into a class of stability representations called predominance diagrams [73], which are common alternatives to the convex hull visualization. Instead of plotting formation energy against composition as is done for the convex hull, these diagrams will instead show the equilibrium phase(s) over a range of intensive thermodynamic inputs, such as chemical potentials of different species. For example, the "chemical potential triangles" shown in Fig. 4 are constructed by first defining ΔE_f in terms of the chemical potentials of each element for all compounds in a chemical space (Ta-Ge-Ir in Fig. 4):

$$\Delta E_{f,A_aX_x\ldots} = \sum_i^{A,X,\ldots} n_i\Delta\mu_i$$

A material is then stable at ($\Delta\mu_A$, $\Delta\mu_X$, …, $\Delta\mu_i$) if and only if no other compound satisfies the condition:

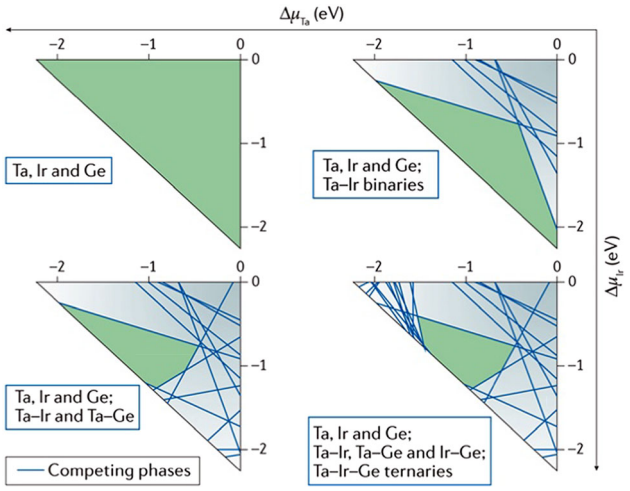

Figure 4 Chemical potential triangles in the Ta-Ir-Ge chemical space to analyze the formation of TaIrGe. The chemical potential of Ta is shown along the x-axis, Ir along the y-axis, and Ge along the diagonal. The set of competing phases considered in each analysis is shown by the blue box in each of the four panels. Blue lines indicate competing phase boundaries in chemical potential space, and the green region indicates the region in chemical potential space where TaIrGe is thermodynamically stable. Reproduced with permission from Ref. [23].

$$\Delta E_{f,A_aX_x\ldots} \leq \sum_i^{A,X,\ldots} n_i\Delta\mu_i$$

In these diagrams, a stable material will appear as a green region in Δμ-space, indicating that this material is on the convex hull ($\Delta E_d \leq 0$) throughout this region. Unstable materials will not appear in predominance diagrams because unlike the convex hull visualization (Fig. 1), there is no energy axis. Because of this, one should be cautious interpreting these diagrams because very small changes in ΔE_f (of the magnitude of ΔE_d) can lead to the appearance or disappearance of certain phases, encouraging the calculation of ΔE_d for all phases in addition to showing their chemical potential window for stability. The effect of including (or excluding) certain competing compounds on the window of stability is illustrated in Fig. 4. Here, as competing phases with increasing components (binary, ternary, etc.) are added to the analysis, the range of chemical potentials over which TaIrGe is stable (the area of the green region) shrinks dramatically, emphasizing the importance of including compounds beyond the standard synthetic precursors when performing stability analyses.

It should be emphasized that these alternative approaches to visualizing stability make use of the same fundamental inputs (ΔE_f and composition) and must agree on the thermodynamic stability of all phases. Indeed, the chemical potential window for the stability of a given phase can also be determined from tangent lines in the convex hull, as shown by the dashed orange line in Fig. 1 (i.e., there is duality between the Gibbs energy or similar thermodynamic potential and chemical potential [73, 74]). Traditionally, chemical potential triangles are more commonly used in the fields of semiconductor design and defect engineering [75], whereas convex hulls have become the standard approach for more general high-throughput materials discovery [76]. Anand et al. recently outlined some of the many connections between these two approaches for visualizing the stability of defects on convex hulls [77]. While both are useful in a variety of applications, in general, predominance diagrams are common when targeting the synthesis of a single phase using synthetic approaches that control chemical potential (e.g., sputtering), whereas convex hulls are common for quickly displaying the thermodynamic stability of a range of phases over some compositional handle(s).

Stability with respect to phase transition

A given chemical composition can crystallize into a wide variety of structures, called polymorphs. Whereas stability with respect to phase separation requires thermodynamic assessment of formation energies as a function of composition, stability with respect to phase transition requires only a comparison of energies at fixed composition. This simplification does not make the problem any easier, however, as the inability to identify the ground-state crystal structure for a given composition has been called "one of the continuing scandals in the physical sciences" [78]. More than 30 years after this declaration, crystal structure prediction (CSP) remains an active area of materials science research. Oganov [79] divides the principal challenges for CSP into two spaces: 1) searching for crystal structures and 2) ranking them by their energies. This work will focus primarily on the second task — determining the stability of a given polymorph with respect to others at fixed composition. The reader is directed to Refs. [79–82] for more focused descriptions of CSP, especially the task of searching for possible crystal structures.

Polymorph energy ordering

Stability with respect to phase transition typically occurs on a smaller energy scale than with respect to phase separation [32]. A convex hull in a chemical space may span a variety of compositions and a few eV/atom in terms of ΔE_f (from highly cohesive compounds with very negative ΔE_f to elemental phases with $\Delta E_f = 0$). In contrast, when the composition is fixed, the energy separation between different polymorphs can be extremely small, sometimes on the order of a few meV/atom. These small energy differences between competing polymorphs emphasize the need to consider higher levels of theory to accurately determine the ground-state structure. For example, Zhang et al. recently showed that for inorganic compounds comprised of main group elements, the SCAN meta-GGA functional systematically outperforms the PBE GGA functional in correctly identifying the ground-state polymorph [83]. Notably, for transition-metal containing compounds, the improvement was not as dramatic, and SCAN still misidentified the ground-state in ~ 15–30% of compounds having polymorphs within 30 meV/atom of the ground-state.

Carefully resolving small polymorph energy differences may require pushing further up Jacob's ladder [84] e.g., with the random-phase approximation (RPA) [85] or quantum Monte Carlo (QMC) [47]. Unlike stability with respect to phase separation, for polymorphic stability, there is much less cancellation of error between competing phases because we are simply comparing the total (and not formation or decomposition) energies of different arrangements of ions at fixed composition. Hence, small changes in ionic positions can lead to subtle energetic effects that, in some cases, can only be captured by sophisticated treatments of the electronic structure. The increased computational expense of these higher-level approaches can be prohibitive for medium- or high-throughput studies, through there has been limited benchmarking on the order of dozens of compounds, at least for ΔE_f [86].

Polymorphic stability under applied conditions

As with the convex hull analysis, stability with respect to phase transition can also be sensitive to the environment. Temperature [87], pressure [88], and substrate [89] (among other factors [90]) can play an important role in the relative energies of competing structures at fixed composition because the arrangement of ions in a structure dictates the response to these applied conditions. For example, we might expect softer polymorphs to generally be preferentially stabilized at high temperature relative to harder polymorphs because of vibrational entropy contributions [53]. Unlike stability with respect to phase separation, however, because composition is fixed for a set of polymorphs, polymorph energies are insensitive to changes in chemical potential.

Solid-state synthesis typically occurs at high temperature to overcome the high activation barriers for self-diffusion in solids. Hence, it is important to understand how the application of temperature affects the relative energies of different polymorphs. Because entropy is the conjugate variable to temperature, entropic contributions can lower the energy of one polymorph relative to the others as a material is heated. As a reference, McCormack and Navrotsky provide a set of known temperature-driven polymorphic phase transitions in oxides in the context of designing high-entropy ceramic materials [91]. Configurational [53] and vibrational [58] entropy usually play the most significant role in contributing to the free energies of crystalline solids [92].

As discussed previously in "Alternative thermodynamic potentials", configurational entropy arises for systems exhibiting disorder. While this is a significant factor that stabilizes disordered alloy systems relative to ordered elemental phases, configurational entropy can only govern polymorph energy ordering if a certain polymorph accommodates disorder more than another (at fixed composition). The phase transition from an ordered intermetallic to a disordered alloy at elevated temperature is a result of configurational entropy lowering the Gibbs free energy of the alloy relative to the intermetallic [93]. A classic example of a configurationally stabilized polymorph outside of intermetallics is the high-temperature superconductor, $YBa_2Cu_3O_{6+x}$ (YBCO), where disorder on the oxygen sublattice in the tetragonal structure contributes to the high-temperature stabilization of this polymorph relative to the ordered orthorhombic phase (Fig. 5). It should be noted that in Fig. 5, vibrational entropy is included in addition to configurational entropy in the determination of the polymorph energy ordering as a function of temperature.

The calculation of vibrational entropy from first principles requires the phonon density of states to be calculated as it is the distribution of these phonon frequencies that give rise to vibrational entropy in solids. As discussed by Stoffel et al., from the calculated phonon density of states, one can obtain the temperature-dependent free phonon energy with Bose–Einstein statistics, and by calculating this quantity at a variety of unit cell volumes, one can bridge the gap between $E(0\ \mathrm{K})$ and $G(T)$ via the quasi-harmonic approximation [94]. In practice, the first-principles calculation of vibrational thermodynamics is a relatively low-throughput procedure, as it requires the often cumbersome calculation of vibrational properties using density functional perturbation theory [95, 96] or the finite displacement method [97, 98] at several cell volumes for each material of interest. Recently, however, the concept of high-throughput phonon calculations has been explored and a couple thousand calculated phonon densities of states are now available in different repositories [99–101]. The vibrational properties of alloy systems have also been discussed in detail by van de Walle and Ceder, who note that the vibrational entropy of formation for alloys can occur on the same scale as configurational entropy ($\sim 0.1 - 0.5$ meV/atom/K) [57]. Although often assumed to have significant effect only at high temperature, there are examples where vibrational entropy distinguishes polymorphs

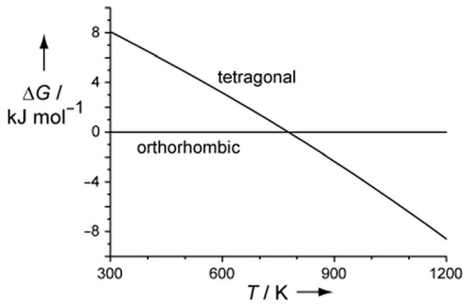

Figure 5 DFT-calculated Gibbs energies of $YBa_2Cu_3O_{6.5}$ relative to the orthorhombic crystal structure. Note that because G is given with respect to an intensive variable, T, $G(T)$ is concave and not convex as it is with respect to composition (extensive). Reproduced with permission from Ref. [94].

even below room temperature, for example, in the hybrid perovskite $CH_3NH_3Mn(N_3)_3$ [102]. At the other end of the temperature spectrum, when materials are near their melting point, the quasi-harmonic approximation loses accuracy as anharmonic effects play a larger role [103, 104].

Dynamic stability

The calculation of the phonon density of states not only assists in determining the vibrational entropy of solid materials but can also reveal the stability of a given crystal structure with respect to energy-lowering collective displacements (i.e., dynamic stability) [52]. When the phonon density of states is computed by either the finite displacement method or density functional perturbation theory, the presence of imaginary modes (negative frequencies) indicates dynamic instability of that crystal structure. The phonon spectrum and presence/absence of imaginary modes also reveals useful information regarding the landscape of potential crystal structures for a given chemical composition. By deforming a dynamically unstable crystal structure along the directions in the Brillouin zone where imaginary modes exist, one can create so-called line diagrams [105] toward dynamically stable polymorphs. A recent example of a line diagram for polymorph exploration in $Bi_2Sn_2O_7$ is shown in Fig. 6 and emphasizes the extremely small energy range that can be important for distinguishing between competing polymorphs [106]. Dynamically stable crystal structures have also been systematically identified using metrics based on order parameters [107] chosen to detect structural stability for chosen subsystems (e.g., binary carbon nitrides) [108]. Note that because the harmonic phonon spectrum is typically calculated at 0 K, imaginary modes can sometimes be removed not only by collective displacements but also by including phonon–phonon interactions in the calculation of temperature-dependent phonon spectra [109]. Anomalies may also arise as a result of approximations in the chosen density functional. For example, certain GGA functionals were shown to give unphysical imaginary phonon modes for PbS and PbTe in their equilibrium structures [110]. Care should be taken to ensure that observed dynamic (in)stabilities are physically meaningful and robust with respect to the chosen calculation procedure.

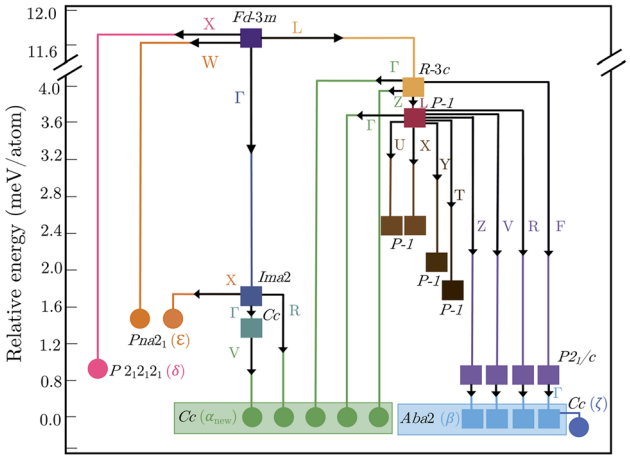

Figure 6 Following phonon modes to map polymorphs in $Bi_2Sn_2O_7$ starting from the high-temperature Fd-3 m phase. Energies are shown relative to the newly discovered lowest energy polymorph, α_{new}. Branches trace the transition pathways obtained by distorting along single imaginary harmonic modes at each step. Circles and squares indicate dynamically stable and unstable structures, respectively, with the space groups as marked. The wavevectors of the imaginary modes connecting pairs of structures are indicated next to the solid lines. Reproduced with permission from Ref. [106].

Ab initio molecular dynamics (AIMD) is a parallel approach sometimes used to probe the dynamic stability of crystalline phases. One approach is to perform AIMD for a candidate structure at some temperature and sample configurations at various time-points to be used as inputs for standard DFT geometry relaxations. If each of these configurations converges to the initial candidate polymorph, this is further (though not conclusive) support for dynamic stability. If, however, lower energy structures are found through this approach, this supports dynamic instability of the initial candidate. This approach is discussed in more detail in the context of 2D materials in Ref. [111]. In a less rigorous approach, one might simply perform an AIMD simulation starting from a candidate polymorph and observe whether the structure changes over the course of the simulation at some (preferably high) temperature. In the author's opinion, this method is not especially conclusive because the transformation of one polymorph into another can require substantial ionic rearrangement that may not occur over the course of a typical AIMD simulation.

Any statements regarding the stability of a given polymorph should account for dynamic stability in addition to stability with respect to phase separation

and transition as discussed previously. These two approaches address stability from fundamentally different perspectives. Dynamic stability addresses the question — once this material forms, will it transform to alternative structures? ΔE_d alternatively addresses the question — can this material form from constituent phases (i.e., be synthesized) at equilibrium? It is common for materials to be dynamically stable, yet still be unstable with respect to decomposition ($\Delta E_d > 0$) into competing compounds, significantly complicating the realization of the desired dynamically stable but non-ground-state phase. Alternatively, if a material is truly dynamically unstable at some conditions, then it should not be possible for this material to be synthesized and persist at those conditions. It could however be possible to remove this dynamic instability by changing the conditions (e.g., by thermal or epitaxial stabilization).

Machine learning for material stability

Predicting the stability (or "formability") of crystalline solids without experiments or electronic structure calculations dates to the early days of solid-state chemistry, with the rules for structural stability by Linus Pauling [112] and the famous perovskite tolerance factor introduced by Viktor Goldschmidt [113]. Fast-forward ~ 100 years and materials scientists now have tremendous databases of experimentally observed crystal structures, DFT-calculated properties for most of these materials (and more), and sufficient computational power to routinely apply state-of-the-art statistical approaches under the umbrella of artificial intelligence and machine learning (ML). Indeed, data-driven analyses of Pauling's rules [114] and Goldschmidt's tolerance factor [115] reveal they may not be as predictive as one would think. Machine learning in materials science is a rapidly growing field, and a comprehensive review is outside the scope of this work. The interested reader is referred to one of many recent reviews on the topic [19–21, 116, 117]. This section will instead focus on the practical challenges of predicting stability with respect to phase separation and phase transition using ML approaches.

Learning stability with respect to phase separation

The direct prediction of ΔE_d with ML is complicated by the non-intrinsic nature of decomposition energies. That is, ΔE_d arises not only from the properties (ΔE_f, composition) of one material but from the neighborhood of competing phases in a particular phase diagram. This also raises questions regarding the completeness of a chemical space. For instance, if an ML model predicts ΔE_d for some material, A_2X_5, in 2021, ΔE_d for that material may change by 2031 as a result of the discovery of new competing phases in the A–X chemical space. Note that this is much more likely to occur with increasing elements in the composition of interest as the ratio of yet-to-be-discovered to already-discovered materials likely increases with chemical complexity (for ternaries, quaternaries, etc.). Despite the non-intrinsic nature of ΔE_d, there have been a few attempts to predict this quantity directly. For example, Schmidt et al., trained a variety of ML models on compositional features (i.e., those that require only the chemical formula and not the structure) on ΔE_d for 20,000 ABX_3 compounds in the cubic perovskite structure, obtaining a mean absolute error (MAE) of 121 meV/atom on an excluded test set [118]. Singstock et al., trained a model to predict ΔE_d for ternary chalcogenides with the chevrel crystal structure and reported an MAE of ~ 30 meV/atom [119]. Ye et al., similarly focused on specific classes of materials, predicting ΔE_d for $C_3A_2D_3O_{12}$ garnets and ABO_3 perovskites and reporting MAEs of ~ 10–30 meV/atom [120]. It should be noted that, in this work, ΔE_d was approximated as the energy of each garnet or perovskite with respect to the linear combination of isovalent binary oxides (e.g., $3CaO + Al_2O_3 + 3SiO_2 \rightarrow Ca_3Al_2Si_3O_{12}$). As discussed previously, the energies of these reactions are lower bounds on ΔE_d because they exclude the potential for ternary and quaternary competing phases, which are prevalent in many chemical spaces.

A more common approach is to train a model on ΔE_f, which is an intrinsic material property, then perform the convex hull construction with predicted ΔE_f (or a mixture of predicted and DFT-calculated ΔE_f) to obtain ΔE_d for each material of interest. There are many examples of both compositional [121–126] and structural [127–131] models to predict ΔE_f, many of which were recently benchmarked on a variety of other material properties [132] in addition to a

focused study on ΔE_d predictions [48]. When comparing the performance of ML models for thermodynamic stability predictions, MAEs were found to be similar for ΔE_f and ΔE_d ($\sim$ 60–150 meV/atom for compositional models; $\sim$ 40 meV/atom for structural models), but MAE does not tell the whole story.

As shown in Fig. 2, the distribution of ΔE_f is much wider than ΔE_d, so the relative error for a fixed MAE is much larger for ΔE_d. Further, it's important to consider how practitioners rely upon thermodynamic stability predictions. Generally, one is interested in identifying the relatively few yet-to-be-discovered stable compounds among many hypothetical candidates. A simulated materials discovery experiment was performed for several compositional models in Ref. [48] by excluding 267 quaternary compounds in the Li-Mn-*M*–O chemical space (*M* = Ti, V, Cr, Fe, Co, Ni, Cu) from training and tasking models with finding the 9 of these compounds that are thermodynamically stable ($\Delta E_d \leq 0$). Of six compositional models studied, none show high correlation between DFT and ML-predicted ΔE_d for this subset of materials, and at most only 3 of the 9 stable materials were successfully identified (Fig. 7). Importantly, when structure is included in the representation (in this case, DFT-optimized structures), CGCNN, was shown to perform much better at this experiment, correctly identifying 5 of 9 stable materials with only 6 false positives [48]. The caveat here is that these predictions also make use of the DFT-optimized crystal structures to make predictions. For real materials discovery problems, predictions will have to be made on unrelaxed or partially relaxed crystal structures. More work is needed to show the efficacy of these models on unrelaxed structures, but a few recent efforts have started to address this problem [133–135].

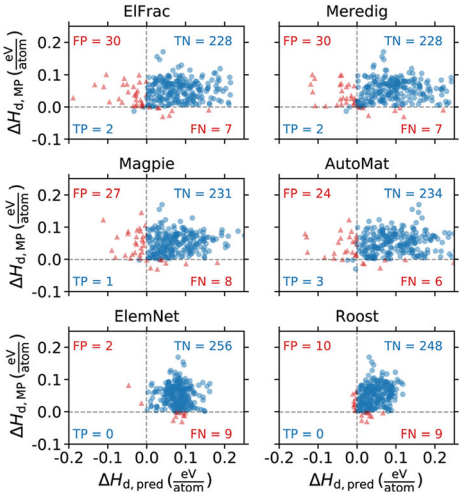

Figure 7 Tasking machine learning models with identifying the few stable compounds among a larger list of quaternary lithium manganese oxides. Actual decomposition energies (from Materials Project) are shown along the *y*-axis with model-predicted decomposition energies for each of six models shown along the *x*-axis. FP = false positive, TP = true positive, TN = true negative, FN = false negative. Reproduced with permission from Ref. [48].

Recent work on thermodynamic stability predictions suggests that general compositional models (those that take as input any arbitrary chemical formula) are incapable of predicting ΔE_d with sufficient accuracy to systematically distinguish between stable and unstable compounds [48]. However, when the condition of generality is softened, and models are developed to focus specifically on a given structure type (e.g., perovskites) or focused on a particular composition (e.g., ABX_3), direct models for ΔE_d (or "formability") or indirect models that predict ΔE_f are expected to perform better [115, 120, 136], though this has not been systematically studied. Additional work in the space of generative machine learning models also supports this proposition, where new low-energy ($\Delta E_d \lesssim 80$ meV/atom) materials were identified once the search space was constrained to the V–O or MgMn–O chemical spaces [137, 138].

Learning stability with respect to phase transition

As discussed previously in the context of Pauling and Goldschmidt, predicting crystal structure from chemical composition is a prevailing challenge in solid-state chemistry. Expanding upon these works (and many more in between), Pettifor developed phenomelogical structure maps that cluster binary compounds crystallizing in the same ground-state structure using the so-called χ-scale [139, 140]. Perhaps unbeknownst to Pettifor at the time of this work in the late 1980s, he was using what we would now refer to as a two-dimensional "descriptor" [141] (χ_A, χ_X for AX compounds) and a very complex decision tree model to determine the boundaries on various crystal structure types (C40, $C11_b$, C54, etc.) for arbitrary A_aX_x compositions. While Pettifor developed these maps using only chemical intuition and hand-drawn decision boundaries, today we make use

of the suite of ML models at our disposal to construct similar structure maps for polymorphic stability. As one example, Ghiringhelli et al. used a compressed sensing approach to not only separate AX compounds into those that crystallize in the rock salt or zinc blende structures, but also predict the energy difference between these two polymorphs within a two-dimensional descriptor (Fig. 8) [142]. The development of descriptor-based approaches for predicting stability with respect to phase transition from chemical composition have been recently aided by developments in symbolic ML [143], where candidate functional expressions are systematically constructed and optimized to perform a learning task [144, 145].

An alternative strategy to symbolic- or descriptor-based approaches is learning structure-energy relationships through ML potentials (MLPs), which are becoming increasingly popular in computational materials science [117, 146–148]. MLPs are typically constructed for a single composition of interest (e.g., TiO_2) and not expected to apply generally to alternative (even if related) compositions (e.g., Ti). This is fundamentally different than general structural representations (CGCNN [128], MEGNet [127], etc.), where training and application is expected to occur over a wide range of both structures and compositions. As such, MLPs are seldom applied to make stability predictions with respect to phase separation (ΔE_d) because these demand predictions over a range of related compositions. They should however be ideally suited to make predictions with respect to phase transition because they are able to learn the subtle effects that configurational changes can have on the energetics of different polymorphs at fixed composition. Indeed, this approach has been used to understand the local energetics governing structure selection in several elemental phases (e.g., boron [149, 150]). Because these potentials are trained to understand local effects, they are also suitable for interrogating non-crystalline phases of these compositions (e.g., amorphous phases [151, 152]).

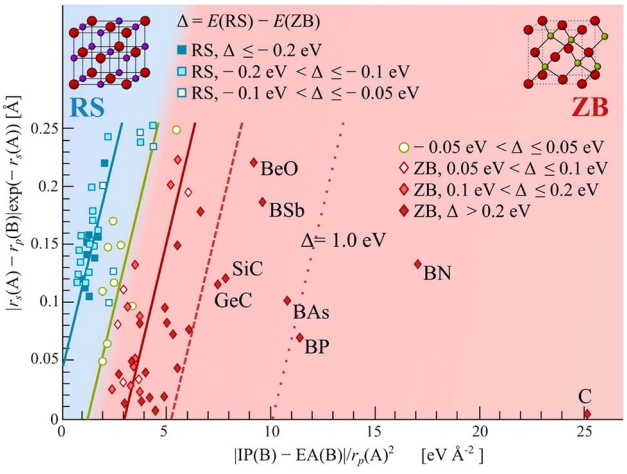

Figure 8 Separating the space of binary compounds into those that crystallize in the rock salt (RS) or zinc blende (ZB) structures. Lines are drawn within each region of the structure map to indicate the energy difference between RS and ZB structures. Reproduced with permission from Ref. [142].

MLPs are also starting to be incorporated with active learning approaches and existing CSP algorithms to accelerate the discovery of new low-energy polymorphs for select compositions. Podryabinkin et al., showed a dramatic acceleration of structure discovery for polymorphs of sodium, boron, and carbon by on-the-fly training of MLPs and integration with evolutionary structure prediction approaches in USPEX [153, 154]. The energetics of various boron clusters were also resolved with increased computational efficiency by Tong et al., by combining MLPs with the CALYPSO structure prediction method [150]. Similar concepts were recently shown to also be tractable for energy ordering in more complex compositions (e.g., Mg_2SiO_4) [155]. Because CSP naturally involves the prediction of novel crystal structures, it is critical to understand the generalizability of any trained model to configurations that do not appear in training. Otherwise, it is likely that these approaches will fail to identify the true ground-state configurations.

Outlook — from stability to synthesizability

Thermodynamic stability is among the most fundamental properties of a material because it provides quantitative information regarding the driving force for forming or decomposing a phase under equilibrium. Ultimately, it is the principal quantity that suggests whether a compound can be formed and if it will persist under conditions of interest. First-principles calculations, open materials databases, and machine learning have proven to be reliable tools to assist in determining thermodynamic stability for many thousands of known and newly proposed materials. However, stability calculations are not the end of the road. There are many metastable materials

($\Delta E_d > 0$) that can be synthesized [32] and many stable phases ($\Delta E_d \leq 0$) for which synthesis can be extremely challenging [156, 157]. Ideally, computational predictions of thermodynamic stability would be directly related to synthesizability, but the complex processes that can occur during synthesis are not purely thermodynamic (i.e., kinetics under difficult-to-describe conditions play a significant role) and today, cannot be easily captured with first-principles calculations [158, 159].

Data-driven surveys of the library of known materials can help us understand, at least generally, the relationships between metastability and synthesizability. For example, Sun et al., showed that the magnitude of accessible metastable phases has a strong dependence on chemistry (e.g., nitrides can be more metastable than oxides) and that the magnitude of metastability is typically larger for phase separation than for phase transition [32]. Aykol et al. showed that the magnitude of accessible polymorphic metastability is also strongly dependent on chemistry, and proposed that polymorphs cannot be synthesized if ΔE_{gs} is larger than the energy of the amorphous phase relative to the ground state [160]. Both of these works shed important insights into the scales of accessible metastability, but it is still difficult to directly translate these insights into reliable "rules of thumb" that indicate the accessibility of a given phase based on ΔE_d or ΔE_{gs}. Thermodynamic calculations of stability are fundamentally limited in terms of predicting the many transformations that can take place during materials synthesis, but they still have proven useful when combined with experimental efforts (often in situ characterization [161]) to understand phase evolution during materials synthesis for specific systems [64, 157, 162–167].

In addition to combined computational and experimental studies of synthesis mechanisms within specific systems, there have also been recent developments toward general synthesis planning and prediction for inorganic materials design. McDermott et al., developed a graph-based approach with a thermodynamically based cost function to identify low-energy routes between specified precursors and target phases [168]. By combining a metric to approximate nucleation kinetics as well as a metric to account for parasitic decomposition reactions, Aykol et al., developed a two-dimensional map where the most likely synthesis reactions for a given target lie on the Pareto frontier [169]. On the topic of polymorphic stability and accessibility, Stevanović et al., developed an approach to map crystal structures to one another and estimate the kinetic barrier to diffusionless transformations between polymorphs [170]. A preeminent challenge in computational materials synthesis prediction is understanding when thermodynamics or kinetics are governing phase evolution and how best to account for the relevant phenomena using calculations. These approaches can help reduce the space of plausible transformations that are likely to take place from a given material or set of materials.

Machine learning may also be a promising avenue for advancing computational predictions of materials synthesis. Jang et al. used partially supervised ML to show there is a predictive relationship between the crystal structure of materials and their synthesizability [171]. Kononova et al. used natural language processing techniques to extract thousands of solid-state synthesis recipes from the published literature, which should prove useful for future data-driven efforts [172]. A major challenge in the prediction of synthesizability is the binary nature of this property — a material is either synthesizable (via some route) or it is not. This is unlike stability, which can be quantified continuously using ΔE_d. Further complicating the matter of synthesis prediction is the nature of scientific research where, with few exceptions, only successful syntheses are currently reported in the literature. These two factors present a major challenge for learning what constitutes synthesizability using data-driven approaches. The burgeoning call to make more data accessible, including the so-called 'failed experiments' [173], will hopefully alleviate this challenge.

Acknowledgements

This work was supported as part of GENESIS: A Next Generation Synthesis Center, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award Number DESC0019212. The author gratefully acknowledges the input of Matthew McDermott and Nathan Szymanski, who provided essential feedback on an early draft of this work. The author would also like to acknowledge Prof. Gerbrand Ceder, Dr. Aaron Holder, Dr. Stephan Lany, Prof. Charles Musgrave, Prof. Vladan Stevanović,

and Prof. Wenhao Sun for extensive discussions on the thermodynamic properties of materials over the last several years.

Declarations

Conflict of interest The author declares no conflict of interest.

References

[1] Jain A, Shin Y, Persson KA (2016) Computational predictions of energy materials using density functional theory. Nat Rev Mater 1(1):15004. https://doi.org/10.1038/natrevmats.2015.4

[2] Alberi K, Nardelli MB, Zakutayev A, Mitas L, Curtarolo S, Jain A, Fornari M, Marzari N, Takeuchi I, Green ML, Kanatzidis M, Toney MF, Butenko S, Meredig B, Lany S, Kattner U, Davydov A, Toberer ES, Stevanovic V, Walsh A, Park N-G, Aspuru-Guzik A, Tabor DP, Nelson J, Murphy J, Setlur A, Gregoire J, Li H, Xiao R, Ludwig A, Martin LW, Rappe AM, Wei S-H, Perkins J (2018) The 2019 materials by design roadmap. J Phys D Appl Phys 52(1):013001. https://doi.org/10.1088/1361-6463/aad926

[3] Shevlin S, Castro B, Li X (2021) Computational materials design. Nat Mater 20(6):727–727. https://doi.org/10.1038/s41563-021-01038-8

[4] Becke AD (2014) Perspective: fifty years of density-functional theory in chemical physics. J Chem Phys 140(18):18A301. https://doi.org/10.1063/1.4869598

[5] Urban A, Seo D-H, Ceder G (2016) Computational understanding of li-ion batteries. npj Comput Mater 2(1):16002. https://doi.org/10.1038/npjcompumats.2016.2

[6] Li Y, Yang K (2021) High-throughput computational design of halide perovskites and beyond for optoelectronics. WIREs Comput Molecular Sci 11(3):e1500. https://doi.org/10.1002/wcms.1500

[7] Gorai P, Stevanović V, Toberer ES (2017) Computationally guided discovery of thermoelectric materials. Nat Rev Mater 2(9):17053. https://doi.org/10.1038/natrevmats.2017.53

[8] Nørskov JK, Bligaard T, Rossmeisl J, Christensen CH (2009) Towards the computational design of solid catalysts. Nat Chem 1(1):37–46. https://doi.org/10.1038/nchem.121

[9] Frey NC, Horton MK, Munro JM, Griffin SM, Persson KA, Shenoy VB (2020) High-throughput search for magnetic and topological order in transition metal oxides. Sci Adv. 6(50):1076. https://doi.org/10.1126/sciadv.abd1076

[10] Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G, Persson KA (2013) Commentary: the materials project: a materials genome approach to accelerating materials innovation. APL Mater 1(1):011002. https://doi.org/10.1063/1.4812323

[11] Curtarolo S, Setyawan W, Hart GLW, Jahnatek M, Chepulskii RV, Taylor RH, Wang S, Xue J, Yang K, Levy O, Mehl MJ, Stokes HT, Demchenko DO, Morgan D (2012) AFLOW: an automatic framework for high-throughput materials discovery. Comput Mater Sci 58:218–226. https://doi.org/10.1016/j.commatsci.2012.02.005

[12] Kirklin S, Saal JE, Meredig B, Thompson A, Doak JW, Aykol M, Rühl S, Wolverton C (2015) The open quantum materials database (OQMD): assessing the accuracy of DFT formation energies. npj Comput Mater 1(1):15010. https://doi.org/10.1038/npjcompumats.2015.10

[13] Pizzi G, Cepellotti A, Sabatini R, Marzari N, Kozinsky B (2016) AiiDA: automated interactive infrastructure and database for computational science. Comput Mater Sci 111:218–230. https://doi.org/10.1016/j.commatsci.2015.09.013

[14] Draxl C, Scheffler M (2018) NOMAD: The FAIR concept for big data-driven materials science. MRS Bull 43(9):676–682. https://doi.org/10.1557/mrs.2018.208

[15] Choudhary K, Garrity KF, Reid ACE, DeCost B, Biacchi AJ, Hight Walker AR, Trautt Z, Hattrick-Simpers J, Kusne AG, Centrone A, Davydov A, Jiang J, Pachter R, Cheon G, Reed E, Agrawal A, Qian X, Sharma V, Zhuang H, Kalinin SV, Sumpter BG, Pilania G, Acar P, Mandal S, Haule K, Vanderbilt D, Rabe K, Tavazza F (2020) The joint automated repository for various integrated simulations (JARVIS) for data-driven materials design. NPJ Comput Mater 6(1):173. https://doi.org/10.1038/s41524-020-00440-1

[16] Horton MK, Dwaraknath S, Persson KA (2021) Promises and perils of computational materials databases. Nature Comput Sci 1(1):3–5. https://doi.org/10.1038/s43588-020-00016-5

[17] Vergniory MG, Elcoro L, Felser C, Regnault N, Bernevig BA, Wang Z (2019) A complete catalogue of high-quality topological materials. Nature 566(7745):480–485. https://doi.org/10.1038/s41586-019-0954-4

[18] Andersen CW, Armiento R, Blokhin E, Conduit GJ, Dwaraknath S, Evans ML, Fekete Á, Gopakumar A, Gražulis S, Merkys A, Mohamed F, Oses C, Pizzi G, Rignanese G-M, Scheidgen M, Talirz L, Toher C, Winston D, Aversa R, Choudhary K, Colinet P, Curtarolo S, Di Stefano D, Draxl C, Er S, Esters M, Fornari M, Giantomassi M, Govoni M, Hautier G, Hegde V, Horton MK, Huck P, Huhs G, Hummelshøj J, Kariryaa A, Kozinsky B, Kumbhar S, Liu M, Marzari N, Morris AJ, Mostofi AA, Persson KA, Petretto G, Purcell T, Ricci F, Rose F, Scheffler M, Speckhard D, Uhrin M, Vaitkus A, Villars P, Waroquiers D,

Wolverton C, Wu M, Yang X (2021) OPTIMADE, an API for exchanging materials data. Sci Data 8(1):217. https://doi.org/10.1038/s41597-021-00974-z

[19] Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A (2018) Machine learning for molecular and materials science. Nature 559(7715):547–555. https://doi.org/10.1038/s41586-018-0337-2

[20] Schmidt J, Marques MRG, Botti S, Marques MAL (2019) Recent advances and applications of machine learning in solid-state materials science. npj Comput Mater 5(1):83. https://doi.org/10.1038/s41524-019-0221-0

[21] Wang AY-T, Murdock RJ, Kauwe SK, Oliynyk AO, Gurlo A, Brgoch J, Persson KA, Sparks TD (2020) Machine learning for materials scientists: an introductory guide toward best practices. Chem Mater 32(12):4954–4965. https://doi.org/10.1021/acs.chemmater.0c01907

[22] Tabor DP, Roch LM, Saikin SK, Kreisbeck C, Sheberla D, Montoya JH, Dwaraknath S, Aykol M, Ortiz C, Tribukait H, Amador-Bedolla C, Brabec CJ, Maruyama B, Persson KA, Aspuru-Guzik A (2018) Accelerating the discovery of materials for clean energy in the era of smart automation. Nat Rev Mater 3(5):5–20. https://doi.org/10.1038/s41578-018-0005-z

[23] Zunger A (2018) Inverse design in search of materials with target functionalities. Nat Rev Chem 2(4):0121. https://doi.org/10.1038/s41570-018-0121

[24] Zunger A (2019) Beware of plausible predictions of fantasy materials. Nature 566(7745):447–449. https://doi.org/10.1038/d41586-019-00676-y

[25] Ong SP, Wang L, Kang B, Ceder G (2008) Li−Fe−P−O2 phase diagram from first principles calculations. Chem Mater 20(5):1798–1807. https://doi.org/10.1021/cm702327g

[26] Bartel CJ, Millican SL, Deml AM, Rumptz JR, Tumas W, Weimer AW, Lany S, Stevanović V, Musgrave CB, Holder AM (2018) Physical descriptor for the gibbs energy of inorganic crystalline solids and temperature-dependent materials chemistry. Nat Commun 9(1):4168

[27] Sun W, Kitchaev DA, Kramer D, Ceder G (2019) Non-equilibrium crystallization pathways of manganese oxides in aqueous solution. Nat Commun 10(1):573. https://doi.org/10.1038/s41467-019-08494-6

[28] Amsler M, Hegde VI, Jacobsen SD, Wolverton C (2018) Exploring the high-pressure materials genome. Phys Rev X 8(4):041021. https://doi.org/10.1103/PhysRevX.8.041021

[29] Sun W, Holder A, Orvañanos B, Arca E, Zakutayev A, Lany S, Ceder G (2017) Thermodynamic routes to novel metastable nitrogen-rich nitrides. Chem Mater 29(16):6936–6946. https://doi.org/10.1021/acs.chemmater.7b02399

[30] Privat R, Jaubert J-N, Berger E, Coniglio L, Lemaitre C, Meimaroglou D, Warth V (2016) Teaching the concept of Gibbs energy minimization through its application to phase-equilibrium calculation. J Chem Educ 93(9):1569–1577. https://doi.org/10.1021/acs.jchemed.6b00205

[31] Bartel CJ, Rumptz JR, Weimer AW, Holder AM, Musgrave CB (2019) High-throughput equilibrium analysis of active materials for solar thermochemical ammonia synthesis. ACS Appl Mater Interfaces 11(28):24850–24858

[32] Sun W, Dacek ST, Ong SP, Hautier G, Jain A, Richards WD, Gamst AC, Persson KA, Ceder G (2016) The thermodynamic scale of inorganic crystalline metastability. Sci Adv 2(11):e1600225. https://doi.org/10.1126/sciadv.1600225

[33] Wang A, Kingsbury R, McDermott M, Horton M, Jain A, Ong SP, Dwaraknath S, Persson KA (2021) A Framework for quantifying uncertainty in DFT energy corrections. Sci Rep 11(1):15496. https://doi.org/10.1038/s41598-021-94550-5

[34] Perdew JP, Burke K, Ernzerhof M (1996) Generalized gradient approximation made simple. Phys Rev Lett 77(18):3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865

[35] Blöchl PE (2000) First-principles calculations of defects in oxygen-deficient silica exposed to hydrogen. Phys Rev B 62(10):6158–6179. https://doi.org/10.1103/PhysRevB.62.6158

[36] Wang L, Maxisch T, Ceder G (2006) Oxidation energies of transition metal oxides within the GGA+U framework. Phys Rev B 73(19):195107. https://doi.org/10.1103/PhysRevB.73.195107

[37] Lany S (2008) Semiconductor thermochemistry in density functional calculations. Phys Rev B 78(24):245207. https://doi.org/10.1103/PhysRevB.78.245207

[38] Stevanović V, Lany S, Zhang X, Zunger A (2012) Correcting density functional theory for accurate predictions of compound enthalpies of formation: fitted elemental-phase reference energies. Phys Rev B 85(11):115104. https://doi.org/10.1103/PhysRevB.85.115104

[39] Kubaschewski O, Kubaschewski O, Alcock CB, Spencer PJ (1993) Materials thermochemistry. International series on materials science and technology. Pergamon Press

[40] Barin I, Sauert F, Schultze-Rhonhof E, Sheng WS (1993) Thermochemical data of pure substances. Thermochemical Data of Pure Substances. VCH

[41] Bale CW, Bélisle E, Chartrand P, Decterov SA, Eriksson G, Gheribi AE, Hack K, Jung I-H, Kang Y-B, Melançon J, Pelton AD, Petersen S, Robelin C, Sangster J, Spencer P, Van Ende M-A (2016) FactSage thermochemical software

and databases, 2010–2016. Calphad 54:35–53. https://doi.org/10.1016/j.calphad.2016.05.002

[42] Chase M (1998) NIST-JANAF thermochemical tables, 4th edition; American Institute of Physics, -1.

[43] Jain A, Hautier G, Ong SP, Moore CJ, Fischer CC, Persson KA, Ceder G (2011) Formation enthalpies by mixing GGA and GGA + U calculations. Phys Rev B 84(4):045115. https://doi.org/10.1103/PhysRevB.84.045115

[44] Sun J, Ruzsinszky A, Perdew JP (2015) Strongly constrained and appropriately normed semilocal density functional. Phys Rev Lett 115(3):036402. https://doi.org/10.1103/PhysRevLett.115.036402

[45] Isaacs EB, Wolverton C (2018) Performance of the strongly constrained and appropriately normed density functional for solid-state materials. Phys Rev Materials 2(6):063801. https://doi.org/10.1103/PhysRevMaterials.2.063801

[46] Bartel CJ, Weimer AW, Lany S, Musgrave CB, Holder AM (2019) The role of decomposition reactions in assessing first-principles predictions of solid stability. npj Comput Mater 5(1):4

[47] Luo Y, Benali A, Shulenburger L, Krogel JT, Heinonen O, Kent PRC (2016) Phase stability of TiO2polymorphs from diffusion quantum monte carlo. New J Phys 18(11):113049. https://doi.org/10.1088/1367-2630/18/11/113049

[48] Bartel CJ, Trewartha A, Wang Q, Dunn A, Jain A, Ceder G (2020) A critical examination of compound stability predictions from machine-learned formation energies. npj Comput Mater 6(1):97. https://doi.org/10.1038/s41524-020-00362-y

[49] Hautier G, Ong SP, Jain A, Moore CJ, Ceder G (2012) 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(15):155208. https://doi.org/10.1103/PhysRevB.85.155208

[50] Goyal A, Gorai P, Peng H, Lany S, Stevanović V (2017) A computational framework for automation of point defect calculations. Comput Mater Sci 130:1–9. https://doi.org/10.1016/j.commatsci.2016.12.040

[51] Hellenbrandt M (2004) The inorganic crystal structure database (ICSD)—present and future. Crystallogr Rev 10(1):17–22. https://doi.org/10.1080/08893110410001664882

[52] Togo A, Tanaka I (2015) First principles phonon calculations in materials science. Scripta Mater 108:1–5. https://doi.org/10.1016/j.scriptamat.2015.07.021

[53] Fultz B (2010) Vibrational thermodynamics of materials. Prog Mater Sci 55(4):247–352. https://doi.org/10.1016/j.pmatsci.2009.05.002

[54] van de Walle A, Asta M, Ceder G (2002) The alloy theoretic automated toolkit: a user guide. Calphad 26(4):539–553. https://doi.org/10.1016/S0364-5916(02)80006-2

[55] Richards WD, Wang Y, Miara LJ, Kim JC, Ceder G (2016) Design of Li1+2xZn1−xPS4, a new lithium ion conductor. Energy Environ Sci 9(10):3272–3278. https://doi.org/10.1039/C6EE02094A

[56] Rost CM, Sachet E, Borman T, Moballegh A, Dickey EC, Hou D, Jones JL, Curtarolo S, Maria J-P (2015) Entropy-stabilized oxides. Nature Commun 6(1):8485. https://doi.org/10.1038/ncomms9485

[57] van de Walle A, Ceder G (2002) The effect of lattice vibrations on substitutional alloy thermodynamics. Rev Mod Phys 74(1):11–45. https://doi.org/10.1103/RevModPhys.74.11

[58] Sutton C, Levchenko SV (2020) First-principles atomistic thermodynamics and configurational entropy. Front Chem 8:757. https://doi.org/10.3389/fchem.2020.00757

[59] Deng Z, Sai Gautam G, Kolli SK, Chotard J-N, Cheetham AK, Masquelier C, Canepa P (2020) Phase behavior in rhombohedral NaSiCON electrolytes and electrodes. Chem Mater 32(18):7908–7920. https://doi.org/10.1021/acs.chemmater.0c02695

[60] Arroyoydedompablo ME, Van der Veaan A, Ceder G (2002) First-principles calculations of lithium ordering and phase stability on Li_xNiO_2. Phys. Rev. B 66(6):064112. https://doi.org/10.1103/PhysRevB.66.064112

[61] Sundman B, Lukas H, Fries S (2007) Computational thermodynamics: the Calphad method. Cambridge University Press, New York

[62] Hillert M (2001) The compound energy formalism. J Alloy Compd 320(2):161–176. https://doi.org/10.1016/S0925-8388(00)01481-X

[63] Liu Z-K (2020) Computational thermodynamics and its applications. Acta Mater 200:745–792. https://doi.org/10.1016/j.actamat.2020.08.008

[64] Miura MA, Bartel CJ, Goto Y, Mizuguchi Y, Moriyoshi C, Kuroiwa Y, Wang Y, Yaguchi T, Shirai M, Nagao M, Rosero-Navarro NC, Tadanaga K, Ceder G, Sun W (2021) Observing and modeling the sequential pairwise reactions that drive solid-state ceramic synthesis. Adv Mater. https://doi.org/10.1002/adma.202100312

[65] Reuter K, Scheffler M (2001) Composition, structure, and stability of RuO_2(110) as a function of oxygen pressure. Phys Rev B 65(3):035406. https://doi.org/10.1103/PhysRevB.65.035406

[66] Canepa P, Dawson JA, Sai Gautam G, Statham JM, Parker SC, Islam MS (2018) Particle morphology and lithium segregation to surfaces of the Li7La3Zr2O12 solid

electrolyte. Chem Mater 30(9):3019–3027. https://doi.org/10.1021/acs.chemmater.8b00649

[67] Shrestha A, Gao X, Hicks JC, Paolucci C (2021) Nanoparticle size effects on phase stability for molybdenum and tungsten carbides. Chem Mater 33(12):4606–4620. https://doi.org/10.1021/acs.chemmater.1c01120

[68] Richards WD, Miara LJ, Wang Y, Kim JC, Ceder G (2016) Interface stability in solid-state batteries. Chem Mater 28(1):266–273. https://doi.org/10.1021/acs.chemmater.5b04082

[69] Xiao Y, Miara LJ, Wang Y, Ceder G (2019) Computational screening of cathode coatings for solid-state batteries. Joule 3(5):1252–1275. https://doi.org/10.1016/j.joule.2019.02.006

[70] Xiao Y, Wang Y, Bo S-H, Kim JC, Miara LJ, Ceder G (2020) Understanding interface stability in solid-state batteries. Nat Rev Mater 5(2):105–126. https://doi.org/10.1038/s41578-019-0157-5

[71] Pourbaix M (1974) Atlas of electrochemical equilibria in aqueous solutions; National Association of Corrosion Engineers

[72] Persson KA, Waldwick B, Lazic P, Ceder G (2012) Prediction of solid-aqueous equilibria: scheme to combine first-principles calculations of solids with experimental aqueous states. Phys Rev B 85(23):235438. https://doi.org/10.1103/PhysRevB.85.235438

[73] Yokokawa H (1999) Generalized chemical potential diagram and its applications to chemical reactions at interfaces between dissimilar materials. J Phase Equilibria 20(3):258. https://doi.org/10.1361/105497199770335794

[74] Sun W, Powell-Palm MJ. Generalized gibbs' phase rule. http://arxiv.org/abs/2105.01337.

[75] Walsh A, Zunger A (2017) Instilling defect tolerance in new compounds. Nat Mater 16(10):964–967. https://doi.org/10.1038/nmat4973

[76] Jain A, Hautier G, Moore CJ, Ping Ong S, Fischer CC, Mueller T, Persson KA, Ceder G (2011) A high-throughput infrastructure for density functional theory calculations. Comput Mater Sci 50(8):2295–2310. https://doi.org/10.1016/j.commatsci.2011.02.023

[77] Anand S, Male JP, Wolverton C, Snyder GJ (2021) Visualizing defect energetics mater. Horiz. https://doi.org/10.1039/D1MH00397F

[78] Maddox J (1988) Crystals from first principles. Nature 335(6187):201–201. https://doi.org/10.1038/335201a0

[79] Oganov AR (2018) Crystal structure prediction: reflections on present status and challenges. Faraday Discuss 211:643–660. https://doi.org/10.1039/C8FD90033G

[80] Oganov AR, Pickard CJ, Zhu Q, Needs RJ (2019) Structure prediction drives materials discovery. Nat Rev Mater 4(5):331–348. https://doi.org/10.1038/s41578-019-0101-8

[81] Woodley SM, Catlow R (2008) Crystal structure prediction from first principles. Nat Mater 7(12):937–946. https://doi.org/10.1038/nmat2321

[82] Pickard CJ, Needs RJ (2011) Ab initiorandom structure searching. J Phys: Condens Matter 23(5):053201. https://doi.org/10.1088/0953-8984/23/5/053201

[83] Zhang Y, Kitchaev DA, Yang J, Chen T, Dacek ST, Sarmiento-Pérez RA, Marques MAL, Peng H, Ceder G, Perdew JP, Sun J (2018) Efficient first-principles prediction of solid stability: towards chemical accuracy. npj Comput Mater 4(1):9. https://doi.org/10.1038/s41524-018-0065-z

[84] Perdew JP, Schmidt K (2001) Jacob's ladder of density functional approximations for the exchange-correlation energy. AIP Conf Proc 577(1):1–20. https://doi.org/10.1063/1.1390175

[85] Zhang M-Y, Cui Z-H, Jiang H (2018) Relative stability of FeS2 polymorphs with the random phase approximation approach. J Mater Chem A 6(15):6606–6616. https://doi.org/10.1039/C8TA00759D

[86] Saritas K, Mueller T, Wagner L, Grossman JC (2017) Investigation of a quantum monte carlo protocol to achieve high accuracy and high-throughput materials formation energies. J Chem Theory Comput 13(5):1943–1951. https://doi.org/10.1021/acs.jctc.6b01179

[87] Grabowski B, Söderlind P, Hickel T, Neugebauer J (2011) Temperature-driven phase transitions from first principles including all relevant excitations: the Fcc-to-Bcc transition in Ca. Phys Rev B 84(21):214107. https://doi.org/10.1103/PhysRevB.84.214107

[88] Wang Y, Ma Y (2014) Perspective: crystal structure prediction at high pressures. J Chem Phys 140(4):040901. https://doi.org/10.1063/1.4861966

[89] Ding H, Dwaraknath SS, Garten L, Ndione P, Ginley D, Persson KA (2016) Computational approach for epitaxial polymorph stabilization through substrate selection. ACS Appl Mater Interf 8(20):13086–13093. https://doi.org/10.1021/acsami.6b01630

[90] Chen B-R, Sun W, Kitchaev DA, Mangum JS, Thampy V, Garten LM, Ginley DS, Gorman BP, Stone KH, Ceder G, Toney MF, Schelhas LT (2018) Understanding crystallization pathways leading to manganese oxide polymorph formation. Nat Commun 9(1):2553. https://doi.org/10.1038/s41467-018-04917-y

[91] McCormack SJ, Navrotsky A (2021) Thermodynamics of high entropy oxides. Acta Mater 202:1–21. https://doi.org/10.1016/j.actamat.2020.10.043

[92] Manzoor A, Pandey S, Chakraborty D, Phillpot SR, Aidhy DS (2018) Entropy contributions to phase stability in binary random solid solutions. npj Computational Materials 4(1):47. https://doi.org/10.1038/s41524-018-0102-y

[93] Ozoliņš V, Wolverton C, Zunger A (1998) Cu-Au, Ag-Au, Cu-Ag, and Ni-Au intermetallics: first-principles study of temperature-composition phase diagrams and structures. Phys Rev B 57(11):6427–6443. https://doi.org/10.1103/PhysRevB.57.6427

[94] Stoffel RP, Wessel C, Lumey M-W, Dronskowski R (2010) Ab initio thermochemistry of solid-state materials. Angew Chem Int Ed 49(31):5242–5266. https://doi.org/10.1002/anie.200906780

[95] Gonze X, Lee C (1997) Dynamical matrices, born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. Phys Rev B – Condens Matter Mater Phys 55(16):10355–10368. https://doi.org/10.1103/PhysRevB.55.10355

[96] Giannozzi P, De Gironcoli S, Pavone P, Baroni S (1991) Ab initio calculation of phonon dispersions in semiconductors. Phys Rev B 43(9):7231–7242. https://doi.org/10.1103/PhysRevB.43.7231

[97] Parlinski K, Li ZQ, Kawazoe Y (1997) First-principles determination of the soft mode in cubic ZrO2. Phys Rev Lett 78(21):4063–4066. https://doi.org/10.1103/PhysRevLett.78.4063

[98] Kresse G, Furthmüller J, Hafner J (1995) Ab initio force constant approach to phonon dispersion relations of diamond and graphite. Europhys Lett (EPL) 32(9):729–734. https://doi.org/10.1209/0295-5075/32/9/005

[99] Phonon database at Kyoto university — phonondb documentation http://phonondb.mtl.kyoto-u.ac.jp/index.html (accessed 2021 -07 -13).

[100] Petretto G, Dwaraknath S, Miranda H, Winston D, Giantomassi M, van Setten MJ, Gonze X, Persson KA, Hautier G, Rignanese G-M (2018) High-throughput density-functional perturbation theory phonons for inorganic materials. Scientific Data 5(1):180065. https://doi.org/10.1038/sdata.2018.65

[101] Talirz L, Kumbhar S, Passaro E, Yakutovich AV, Granata V, Gargiulo F, Borelli M, Uhrin M, Huber SP, Zoupanos S, Adorf CS, Andersen CW, Schütt O, Pignedoli CA, Passerone D, VandeVondele J, Schulthess TC, Smit B, Pizzi G, Marzari N (2020) Materials cloud, a platform for open computational science. Sci Data 7(1):299. https://doi.org/10.1038/s41597-020-00637-5

[102] Wei W, Li W, Butler KT, Feng G, Howard CJ, Carpenter MA, Lu P, Walsh A, Cheetham AK (2018) An unusual phase transition driven by vibrational entropy changes in a hybrid organic-inorganic perovskite. Angew Chem Int Ed 57(29):8932–8936. https://doi.org/10.1002/anie.201803176

[103] Grabowski B, Ismer L, Hickel T, Neugebauer J (2009) Ab Initio up to the melting point: anharmonicity and vacancies in aluminum. Phys Rev B - Condensed Matter Mater Phys. https://doi.org/10.1103/PhysRevB.79.134106

[104] Hellman O, Abrikosov IA, Simak SI (2011) Lattice dynamics of anharmonic solids from first principles. Phys Rev B 84(18):180301. https://doi.org/10.1103/PhysRevB.84.180301

[105] Togo A, Tanaka I (2013) Evolution of crystal structures in metallic elements. Phys Rev B – Condens Matter Mater Phys. https://doi.org/10.1103/PhysRevB.87.184104

[106] Rahim W, Skelton JM, Savory CN, Evans IR, Evans JSO, Walsh A, Scanlon DO (2020) Polymorph exploration of bismuth stannate using first-principles phonon mode mapping. Chem Sci 11(30):7904–7909. https://doi.org/10.1039/D0SC02995E

[107] Zimmermann NER, Jain A (2020) Local structure order parameters and site fingerprints for quantification of coordination environment and crystal structure similarity. RSC Adv 10(10):6063–6081. https://doi.org/10.1039/C9RA07755C

[108] Villarreal R, Singh P, Arroyave R (2021) Metric-driven search for structurally stable inorganic compounds. Acta Mater 202:437–447. https://doi.org/10.1016/j.actamat.2020.10.055

[109] Zhao X-G, Yang D, Sun Y, Li T, Zhang L, Yu L, Zunger A (2017) Cu–In Halide Perovskite Solar Absorbers. J. Am. Chem. Soc. 139(19):6718–6725. https://doi.org/10.1021/jacs.7b02120

[110] Skelton JM, Tiana D, Parker SC, Togo A, Tanaka I, Walsh A (2015) Influence of the exchange-correlation functional on the quasi-harmonic lattice dynamics of II-VI semiconductors. J Chem Phys 143(6):064710. https://doi.org/10.1063/1.4928058

[111] Malyi OI, Sopiha KV, Persson C (2019) Energy, phonon, and dynamic stability criteria of two-dimensional materials. ACS Appl Mater Interfaces 11(28):24876–24884. https://doi.org/10.1021/acsami.9b01261

[112] Pauling L (1929) The principles determining the structure of complex ionic crystals. J Am Chem Soc 51(4):1010–1026. https://doi.org/10.1021/ja01379a006

[113] Goldschmidt VM (1929) Crystal structure and chemical constitution. Trans Faraday Soc 25:253–283. https://doi.org/10.1039/TF9292500253

[114] George J, Waroquiers D, Di Stefano D, Petretto G, Rignanese G-M, Hautier G (2020) The limited predictive power of the pauling rules. Angew Chem Int Ed 59(19):7569–7575. https://doi.org/10.1002/anie.202000829

[115] Bartel CJ, Sutton C, Goldsmith BR, Ouyang R, Musgrave CB, Ghiringhelli LM, Scheffler M (2019) New tolerance factor to predict the stability of perovskite oxides and halides. Sci Adv 5(2):0693

[116] Schleder GR, Padilha ACM, Acosta CM, Costa M, Fazzio A (2019) From DFT to machine learning: recent approaches to materials science–a review. J Phys Mater 2(3):032001. https://doi.org/10.1088/2515-7639/ab084b

[117] Miksch AM, Morawietz T, Kästner J, Urban A, Artrith N (2021) Strategies for the construction of machine-learning potentials for accurate and efficient atomic-scale simulations. Mach Learn Sci Technol 2(3):031001. https://doi.org/10.1088/2632-2153/abfd96

[118] Schmidt J, Shi J, Borlido P, Chen L, Botti S, Marques MAL (2017) Predicting the thermodynamic stability of solids combining density functional theory and machine learning. Chem Mater 29(12):5090–5103. https://doi.org/10.1021/acs.chemmater.7b00156

[119] Singstock NR, Ortiz-Rodríguez JC, Perryman JT, Sutton C, Velázquez JM, Musgrave CB (2021) Machine learning guided synthesis of multinary chevrel phase chalcogenides. J Am Chem Soc 143(24):9113–9122. https://doi.org/10.1021/jacs.1c02971

[120] Ye W, Chen C, Wang Z, Chu I-H, Ong SP (2018) Deep neural networks for accurate predictions of crystal stability. Nat Commun 9(1):3800. https://doi.org/10.1038/s41467-018-06322-x

[121] Ward L, Agrawal A, Choudhary A, Wolverton C (2016) A general-purpose machine learning framework for predicting properties of inorganic materials. npj Comput Mater 2(1):16028. https://doi.org/10.1038/npjcompumats.2016.28

[122] Meredig B, Agrawal A, Kirklin S, Saal JE, Doak JW, Thompson A, Zhang K, Choudhary A, Wolverton C (2014) Combinatorial screening for new materials in unconstrained composition space with machine learning. Phys Rev B 89(9):094104. https://doi.org/10.1103/PhysRevB.89.094104

[123] Jha D, Ward L, Paul A, Liao W, Choudhary A, Wolverton C, Agrawal A (2018) ElemNet: deep learning the chemistry of materials from only elemental composition. Sci Rep 8(1):17593. https://doi.org/10.1038/s41598-018-35934-y

[124] Wang AY-T, Kauwe SK, Murdock RJ, Sparks TD (2021) Compositionally restricted attention-based network for materials property predictions. npj Computational Materials 7(1):77. https://doi.org/10.1038/s41524-021-00545-1

[125] Goodall REA, Lee AA (2020) Predicting materials properties without crystal structure: deep representation learning from stoichiometry. Nat Commun 11(1):6280. https://doi.org/10.1038/s41467-020-19964-7

[126] Peterson G, Brgoch J (2021) Materials discovery through machine learning formation energy. J Phys Energy. https://doi.org/10.1088/2515-7655/abe425

[127] Chen C, Ye W, Zuo Y, Zheng C, Ong SP (2019) Graph networks as a universal machine learning framework for molecules and crystals. Chem Mater 31(9):3564–3572. https://doi.org/10.1021/acs.chemmater.9b01294

[128] Xie T, Grossman JC (2018) Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. Phys Rev Lett 120(14):145301. https://doi.org/10.1103/PhysRevLett.120.145301

[129] Schütt KT, Sauceda HE, Kindermans P-J, Tkatchenko A, Müller K-R (2018) SchNet – a deep learning architecture for molecules and materials. J Chem Phys 148(24):241722. https://doi.org/10.1063/1.5019779

[130] Park CW, Wolverton C (2020) Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery. Phys Rev Mater 4(6):063801. https://doi.org/10.1103/PhysRevMaterials.4.063801

[131] Faber FA, Lindmaa A, von Lilienfeld OA, Armiento R (2016) Machine learning energies of 2 million Elpasolite ABC2D6 crystals. Phys Rev Lett 117(13):135502. https://doi.org/10.1103/PhysRevLett.117.135502

[132] Dunn A, Wang Q, Ganose A, Dopp D, Jain A (2020) Benchmarking materials property prediction methods: the matbench test set and automatminer reference algorithm. npj Comput Mater 6(1):138. https://doi.org/10.1038/s41524-020-00406-3

[133] Pandey S, Qu J, Stevanovic V, John PS, Gorai P (2021) A graph neural network for predicting energy and stability of known and hypothetical crystal structures.https://doi.org/10.26434/chemrxiv.14428865.v1

[134] Zuo Y, Qin M, Chen C, Ye W, Li X, Luo J, Ong SP (2021) Accelerating materials discovery with bayesian optimization and graph deep learning. [cond-mat].http://arxiv.org/abs/2104.10242

[135] Goodall REA, Parackal AS, Faber FA, Armiento R, Lee AA (2021) Rapid discovery of novel materials by coordinate-free coarse graining. [cond-mat, physics:physics].http://arxiv.org/abs/2106.11132

[136] Balachandran PV, Emery AA, Gubernatis JE, Lookman T, Wolverton C, Zunger A (2018) Predictions of new ABO_3 perovskite compounds by combining machine learning and density functional theory. Phys Rev Mater 2(4):043802. https://doi.org/10.1103/PhysRevMaterials.2.043802

[137] Noh J, Kim J, Stein HS, Sanchez-Lengeling B, Gregoire JM, Aspuru-Guzik A, Jung Y (2019) Inverse design of solid-state materials via a continuous representation. Matter

1(5):1370–1384. https://doi.org/10.1016/j.matt.2019.08.017

[138] Kim S, Noh J, Gu GH, Aspuru-Guzik A, Jung Y (2020) Generative adversarial networks for crystal structure prediction. ACS Cent Sci 6(8):1412–1420. https://doi.org/10.1021/acscentsci.0c00426

[139] Pettifor DG (1986) The structures of binary compounds. I. Phenomenological structure maps. J Phys C: Solid State Phys 19(3):285–313. https://doi.org/10.1088/0022-3719/19/3/002

[140] Pettifor DG (1984) A chemical scale for crystal-structure maps. Solid State Commun 51(1):31–34. https://doi.org/10.1016/0038-1098(84)90765-8

[141] Ghiringhelli LM, Vybiral J, Levchenko SV, Draxl C, Scheffler M (2015) Big data of materials science: critical role of the descriptor. Phys Rev Lett 114(10):105503. https://doi.org/10.1103/PhysRevLett.114.105503

[142] Ghiringhelli LM, Vybiral J, Ahmetcik E, Ouyang R, Levchenko SV, Draxl C, Scheffler M (2017) Learning physical descriptors for materials science by compressed sensing. New J Phys 19(2):023017. https://doi.org/10.1088/1367-2630/aa57bf

[143] Wang Y, Wagner N, Rondinelli JM (2019) Symbolic regression in materials science. MRS Commun 9(3):793–805. https://doi.org/10.1557/mrc.2019.85

[144] Ouyang R, Curtarolo S, Ahmetcik E, Scheffler M, Ghiringhelli LM (2018) SISSO: a compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. Phys Rev Materials 2(8):083802. https://doi.org/10.1103/PhysRevMaterials.2.083802

[145] Ouyang R, Ahmetcik E, Carbogno C, Scheffler M, Ghiringhelli LM (2019) Simultaneous learning of several materials properties from incomplete databases with multi-task SISSO. J Phys Mater 2(2):024002. https://doi.org/10.1088/2515-7639/ab077b

[146] Behler J (2016) Perspective: machine learning potentials for atomistic simulations. J Chem Phys 145(17):170901. https://doi.org/10.1063/1.4966192

[147] Deringer VL, Caro MA, Csányi G (2019) Machine learning interatomic potentials as emerging tools for materials science. Adv Mater 31(46):1902765. https://doi.org/10.1002/adma.201902765

[148] Tong Q, Gao P, Liu H, Xie Y, Lv J, Wang Y, Zhao J (2020) Combining machine learning potential and structure prediction for accelerated materials design and discovery. J Phys Chem Lett 11(20):8710–8720. https://doi.org/10.1021/acs.jpclett.0c02357

[149] Deringer VL, Pickard CJ, Csányi G (2018) Data-driven learning of total and local energies in elemental Boron. Phys Rev Lett 120(15):156001. https://doi.org/10.1103/PhysRevLett.120.156001

[150] Tong Q, Xue L, Lv J, Wang Y, Ma Y (2018) Accelerating CALYPSO structure prediction by data-driven learning of a potential energy surface. Faraday Discuss 211:31–43. https://doi.org/10.1039/C8FD00055G

[151] Mocanu FC, Konstantinou K, Lee TH, Bernstein N, Deringer VL, Csányi G, Elliott SR (2018) Modeling the phase-change memory material, Ge2Sb2Te5, with a machine-learned interatomic potential. J Phys Chem B 122(38):8998–9006. https://doi.org/10.1021/acs.jpcb.8b06476

[152] Deringer VL, Bernstein N, Csányi G, Ben Mahmoud C, Ceriotti M, Wilson M, Drabold DA, Elliott SR (2021) Origins of structural and electronic transitions in disordered silicon. Nature 589(7840):59–64. https://doi.org/10.1038/s41586-020-03072-z

[153] Glass CW, Oganov AR, Hansen N (2006) USPEX—evolutionary crystal structure prediction. Comput Phys Commun 175:713–720. https://doi.org/10.1016/j.cpc.2006.07.020

[154] Podryabinkin EV, Tikhonov EV, Shapeev AV, Oganov AR (2019) Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. Phys Rev B 99(6):064114. https://doi.org/10.1103/PhysRevB.99.064114

[155] Hong C, Choi JM, Jeong W, Kang S, Ju S, Lee K, Jung J, Youn Y, Han S (2020) Training machine-learning potentials for crystal structure prediction using disordered structures. Phys Rev B 102(22):224104. https://doi.org/10.1103/PhysRevB.102.224104

[156] Wustrow A, Key B, Phillips PJ, Sa N, Lipton AS, Klie RF, Vaughey JT, Poeppelmeier KR (2018) Synthesis and characterization of MgCr2S4 thiospinel as a potential magnesium cathode. Inorg Chem 57(14):8634–8638. https://doi.org/10.1021/acs.inorgchem.8b01417

[157] Miura A, Ito H, Bartel CJ, Sun W, Rosero-Navarro NC, Tadanaga K, Nakata H, Maeda K, Ceder G (2020) Selective metathesis synthesis of MgCr2S4 by control of thermodynamic driving forces. Mater Horiz 7(5):1310–1316. https://doi.org/10.1039/C9MH01999E

[158] Chamorro JR, McQueen TM (2018) Progress toward solid state synthesis by design. Acc Chem Res 51(11):2918–2925. https://doi.org/10.1021/acs.accounts.8b00382

[159] Kovnir K (2021) Predictive synthesis. Chem Mater 33(13):4835–4841. https://doi.org/10.1021/acs.chemmater.1c01484

[160] Aykol M, Dwaraknath SS, Sun W, Persson KA (2018) Thermodynamic limit for synthesis of metastable inorganic

materials. Sci Adv 4(4):0148. https://doi.org/10.1126/sciadv.aaq0148

[161] Shoemaker DP, Hu Y-J, Chung DY, Halder GJ, Chupas PJ, Soderholm L, Mitchell JF, Kanatzidis MG (2014) In situ studies of a platform for metastable inorganic crystal growth and materials discovery. Proc Natl Acad Sci USA 111(30):10922. https://doi.org/10.1073/pnas.1406211111

[162] Bianchini M, Wang J, Clément RJ, Ouyang B, Xiao P, Kitchaev D, Shi T, Zhang Y, Wang Y, Kim H, Zhang M, Bai J, Wang F, Sun W, Ceder G (2020) The interplay between thermodynamics and kinetics in the solid-state synthesis of layered oxides. Nat Mater 19(10):1088–1095. https://doi.org/10.1038/s41563-020-0688-6

[163] McClain R, Malliakas CD, Shen J, Wolverton C, Kanatzidis MG (2021) In situ mechanistic studies of two divergent synthesis routes forming the heteroanionic BiOCuSe. J Am Chem Soc. https://doi.org/10.1021/jacs.1c03947

[164] Todd PK, Wustrow A, McAuliffe RD, McDermott MJ, Tran GT, McBride BC, Boeding ED, O'Nolan D, Liu C-H, Dwaraknath SS, Chapman KW, Billinge SJL, Persson KA, Huq A, Veith GM, Neilson JR (2020) Defect-Accommodating intermediates yield selective low-temperature synthesis of YMnO3 polymorphs. Inorg Chem 59(18):13639–13650. https://doi.org/10.1021/acs.inorgchem.0c02023

[165] Rom CL, Fallon MJ, Wustrow A, Prieto AL, Neilson JR (2021) Bulk synthesis, structure, and electronic properties of magnesium zirconium nitride solid solutions. Chem Mater 33(13):5345–5354. https://doi.org/10.1021/acs.chemmater.1c01450

[166] He H, Yee C-H, McNally DE, Simonson JW, Zellman S, Klemm M, Kamenov P, Geschwind G, Zebro A, Ghose S, Bai J, Dooryhee E, Kotliar G, Aronson MC (2018) Combined computational and experimental investigation of the $La_2CuO_{4-x}S_{x}$ (0 ≤ x</Em> ≤ 4) quaternary system. Proc Natl Acad Sci USA 115(31):7890. https://doi.org/10.1073/pnas.1800284115

[167] Todd PK, McDermott MJ, Rom CL, Corrao AA, Denney JJ, Dwaraknath SS, Khalifah PG, Persson KA, Neilson JR (2021) Selectivity in Yttrium manganese oxide synthesis via local chemical potentials in hyperdimensional phase space. J Am Chem Soc 143(37):15185–15194. https://doi.org/10.1021/jacs.1c06229

[168] McDermott MJ, Dwaraknath SS, Persson KA (2021) A graph-based network for predicting chemical reaction pathways in solid-state materials synthesis. Nat Commun 12(1):3097. https://doi.org/10.1038/s41467-021-23339-x

[169] Aykol M, Montoya JH, Hummelshøj J (2021) Rational solid-state synthesis routes for inorganic materials. J Am Chem Soc 143(24):9244–9259. https://doi.org/10.1021/jacs.1c04888

[170] Stevanović V, Trottier R, Musgrave C, Therrien F, Holder A, Graf P (2018) Predicting kinetics of polymorphic transformations from structure mapping and coordination analysis. Phys Rev Materials 2(3):033802. https://doi.org/10.1103/PhysRevMaterials.2.033802

[171] Jang J, Gu GH, Noh J, Kim J, Jung Y (2020) Structure-based synthesizability prediction of crystals using partially supervised learning. J Am Chem Soc 142(44):18836–18843. https://doi.org/10.1021/jacs.0c07384

[172] Kononova O, Huo H, He T, Rong Z, Botari T, Sun W, Tshitoyan V, Ceder G (2019) Text-mined dataset of inorganic materials synthesis recipes. Scientific Data 6(1):203. https://doi.org/10.1038/s41597-019-0224-1

[173] Raccuglia P, Elbert KC, Adler PDF, Falk C, Wenny MB, Mollo A, Zeller M, Friedler SA, Schrier J, Norquist AJ (2016) Machine-learning-assisted materials discovery using failed experiments. Nature 533(7601):73–76. https://doi.org/10.1038/nature17439