

MATERIALS SCIENCE

The thermodynamic scale of inorganic crystalline metastability

Wenhao Sun,^{1,2} Stephen T. Dacek,¹ Shyue Ping Ong,³ Geoffroy Hautier,⁴ Anubhav Jain,⁵ William D. Richards,¹ Anthony C. Gamst,⁶ Kristin A. Persson,^{5,7} Gerbrand Ceder^{1,2,7*}

The space of metastable materials offers promising new design opportunities for next-generation technological materials such as complex oxides, semiconductors, pharmaceuticals, steels, and beyond. Although metastable phases are ubiquitous in both nature and technology, only a heuristic understanding of their underlying thermodynamics exists. We report a large-scale data-mining study of the Materials Project, a high-throughput database of density functional theory–calculated energetics of Inorganic Crystal Structure Database structures, to explicitly quantify the thermodynamic scale of metastability for 29,902 observed inorganic crystalline phases. We reveal the influence of chemistry and composition on the accessible thermodynamic range of crystalline metastability for polymorphic and phase-separating compounds, yielding new physical insights that can guide the design of novel metastable materials. We further assert that not all low-energy metastable compounds can necessarily be synthesized, and propose a principle of ‘remnant metastability’—that observable metastable crystalline phases are generally remnants of thermodynamic conditions where they were once the lowest free-energy phase.

INTRODUCTION

Since the formulation of materials thermodynamics by Josiah Willard Gibbs in 1878 (1), a major paradigm in materials science and engineering has been the identification and synthesis of thermodynamically stable materials. However, metastable phases—kinetically trapped phases with positive free energy above the equilibrium state—are ubiquitous in nature, industry, and the laboratory. For numerous materials technologies, metastable phases can exhibit superior properties than their corresponding stable phases; examples can be found for photocatalysts (2), photovoltaics (3, 4), gas sorbents (5), ion conductors (6), pharmaceuticals (7), steels (8), and more. This suggests that the oft-overlooked space of metastable phases may yield promising new technological materials. Unfortunately, principles for the design and synthesis of inorganic metastable phases remain largely heuristic in nature (9, 10). Crucially lacking is a fundamental understanding of which metastable phases can be synthesized, and whether or not synthesizability is related to the excess enthalpy of a metastable phase above its thermodynamic ground state (11, 12). Answering this would enable a more rational approach toward the design of novel metastable materials.

To better predict which metastable materials can be made, we first seek to understand the metastable materials that have been made. In this work, we investigate the thermodynamic scale of observed metastable materials, quantified by the difference in enthalpy between metastable compounds and their corresponding ground-state phase(s). This thermodynamic analysis provides a baseline understanding of crystalline metastability, and can be a foundation on which future kinetic theories—involving transformation barriers and metastable lifetimes—can be constructed. In practice, the magnitude of this energy

scale is difficult to survey comprehensively, as the scope of materials that exhibit metastability is enormous—spanning the allotropes of pure elements, metal alloys, minerals, semiconductors, ceramics, salts, and more. Furthermore, experimental calorimetry of metastable phases is available for only a limited subset of these. Computation can probe the complex energy landscape of competing phases in specific chemical systems (13–16), but these studies generally lack the scope to make broader statements about how the energy scale of metastability varies by chemistry and composition.

The recent advent of high-throughput computational materials science has enabled the construction of materials property databases with the potential to span all known inorganic materials, providing new opportunities to conduct large-scale materials data analyses (17–20). Here, we explicitly quantify the thermodynamic scale of inorganic crystalline metastability by data-mining the Materials Project database (18), which is composed of materials properties obtained by density functional theory (DFT) calculations performed on crystal structures obtained from the Inorganic Crystal Structure Database (ICSD) (21). This strategy leverages the fact that the crystal structures of metastable phases have been more thoroughly characterized than their thermodynamic properties, so by evaluating the energies of materials phases computationally, we can survey nearly all known inorganic materials within a consistent thermodynamic framework.

Phase stability analysis and data curation

The Materials Project uses extensive tools to compute first-principles phase stability across multinary spaces (22, 23), with DFT +*U* corrections for strongly correlated compounds (24), and gas-phase chemical potentials (N₂, O₂, etc.) fit from experimental decomposition energies (25). A detailed discussion of the high-throughput methodology and infrastructure of the Materials Project can be found in the study by Jain *et al.* (26). We have previously benchmarked DFT to accurately predict ground-state phases over 90% of the time (27, 28) and that formation energies from adjacent stable compounds in phase space can be accurately calculated to within 24 meV/atom (29). However, we emphasize that errors in the DFT formation energy (30, 31) cannot be directly applied as error bars in phase stability and metastability, which would require convex hull analysis.

¹Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 USA. ²Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. ³Department of NanoEngineering, University of California, San Diego, La Jolla, CA 92093, USA. ⁴Institute of Condensed Matter and Nanosciences, Université catholique de Louvain, Louvain-la-Neuve 1348, Belgium. ⁵Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. ⁶Computational and Applied Statistics Laboratory, Department of Mathematics, University of California, San Diego, La Jolla, CA 92093, USA. ⁷Department of Materials Science and Engineering, University of California, Berkeley, CA 94720, USA.

*Corresponding author. Email: gceder@berkeley.edu

We perform this analysis by bootstrapping Monte Carlo simulations of convex hulls jittered by random DFT formation energy error of 24 meV/atom (see section S2.1). We construct standard deviations on the overall percentage of metastable phases and 95% confidence intervals on the energy scale of metastability. For the average convex hull, which contains 65 entries, we obtain a 4% standard deviation in the fraction of metastable compounds and a 28% probability of obtaining different stable and metastable compositions. These are likely overestimations of DFT phase stability errors, as we use random DFT formation energy error, whereas the systematic nature of DFT should introduce correlations into formation energy errors within a given chemical space, reducing the magnitude of error in computed metastability and phase stability. We also apply order statistics to obtain 95% confidence intervals on the reported medians and 90th percentiles of the thermodynamic scale of metastability (sections S3.1 and S3.2). In general, the calculated error bars and confidence intervals are much smaller in magnitude than the observed differences across chemistries and compositions, allowing us to conclude that the trends we observe between categories are statistically significant.

Not all ICSD entries are relevant to this study—we are only interested in the subset of ICSD entries corresponding to observed, bulk crystalline phases, and whose energies are well-described by DFT. Great care was taken to curate our data set to satisfy these criteria. We first filtered out ICSD entries that are poorly described for DFT, notably those with missing light atoms, disorder, or van der Waals bonding. We also removed redundant and duplicate entries, as it would result in double counting, and thus overrepresentation of well-studied materials. The enormous size of the ICSD necessitated the use of automatic screening algorithms for these two steps, by which we were able to reduce the 150,000+ entries of the 2012 ICSD to 29,902 unique phases. There are two additional classes of ICSD entries that are spurious to our investigation: defect unit cells, which form unphysical bulk crystals when repeated by DFT periodic boundary conditions, and unobserved hypothetical, computationally-constructed structures. We manually surveyed the frequency and energy distribution of these entries, and found that they compose approximately 20% of the ICSD and represent the highest-energy calculated phases in the database (section S2.4). We avoid these entries by conducting our statistical analyses on the 80% lowest-energy metastable compounds per query, which is a coarse but necessary strategy to avoid this spurious data in our statistical analyses. A detailed discussion of data provenance and validation can be found in sections S2.2 to S2.4.

RESULTS

The thermodynamic scale of inorganic crystalline metastability

We begin our analysis by defining the thermodynamic metastability of a compound to be its zero pressure, $T = 0$ K enthalpy above the ground-state phase(s), which for polymorphs is the lowest-energy compound of the same composition and for phase-separating materials is the linear combination of energies of the stable phase-separated decomposition products. Under applied thermodynamic conditions (temperature, pressure, etc.), the relative stability between phases may change, but the collective thermodynamic scale of metastability unlikely deviates significantly from the one surveyed at zero temperature and pressure. We describe higher excess-enthalpy phases as “more” metastable and express metastability in units of millielec-

tron volts per atom (10 meV/atom \sim 1 kJ/mol atoms), normalized per atom (not per formula unit).

Of the 29,902 provenance-filtered Materials Project entries, $50.5 \pm 4\%$ (15,097) are metastable, with an approximately exponentially decreasing probability distribution of metastability versus frequency (Fig. 1A). The DFT-calculated median metastability of all known inorganic crystalline materials is 15 ± 0.5 meV/atom, and the 90th percentile is 67 ± 2 meV/atom. We next analyze metastability grouped by chemistry, number of elements in a compound, and phase-separating versus polymorphic metastability. However, readers interested in a more involved investigation of specific systems can readily reconstruct our complete data set from the Materials Project for free, via its application programming interface (API) (32) using the procedures described in section S2.2.

Influence of chemistry

We first group the entries by chemistry, defined by the most electronegative element in a compound. In general, we observe that the stronger the average cohesive energy for a given chemistry, the greater the accessible crystalline metastability. In Fig. 1B, we illustrate this observation for the group VI chemistries, where we map the bivariate sample density as a function of metastability and cohesive energy, where darker regions indicate higher scatter density. There is a clear trend that of the group VI chemistries, oxides exhibit the strongest average cohesive energy and the largest range of metastability, and as one descends down the periodic table, diminished average lattice cohesivity yields smaller accessible energy ranges of metastability. This observation seems in line with conventional wisdom that stronger cohesion and bonding can stabilize higher-energy atomic arrangements, which allows for thermodynamically metastable compounds to resist transformation to the ground state.

We next broaden our observation to all mixed ionic/covalent solids by including the group V and group VII chemistries. Because the probability distribution of metastability has a similar shape across all chemistries (Fig. 1A), we characterize the metastability of a chemistry class by its median and 90th percentile. In Fig. 1C, these medians and 90th percentiles are represented as vertices of line segments, arranged vertically by the median cohesive energy for each chemistry. Between periodic groups, we observe that the cohesive energy becomes stronger with greater anionic charge—that is, in the order (group VII)⁻ < (group VI)²⁻ < (group V)³⁻—reflecting the significance of the electrostatic contribution to the cohesive energy. Within each group, we confirm the general observation that the stronger the cohesive energy of a chemistry, the higher the accessible metastability, with fluorides, oxides, and nitrides exhibiting the greatest energy scales of metastability in their respective groups.

It follows from our arguments that nitrides should exhibit the highest energy scale of metastability, which we observe—but the magnitude is somewhat surprising, with a median at 63 meV/atom and the 90th percentile at an exceptional 195 meV/atom. Nitrides tend to form very strong ionic/covalent bonds and are the only chemistry capable of forming triple bonds in the solid state. In addition to the great diversity of known nitride structures (33, 34), structural investigations of the high-energy nitrides do reveal many structures with strongly directional covalent bonds (for example, boron and carbon nitrides), as well as triple-bonded cyanide anions [such as Hg(CN)₂ and FeNi₂(CN)₆]. These strong bonds can lock in energetically unfavorable atomic arrangements, enabling the persistence of highly metastable (>70 meV/atom) nitride structures.

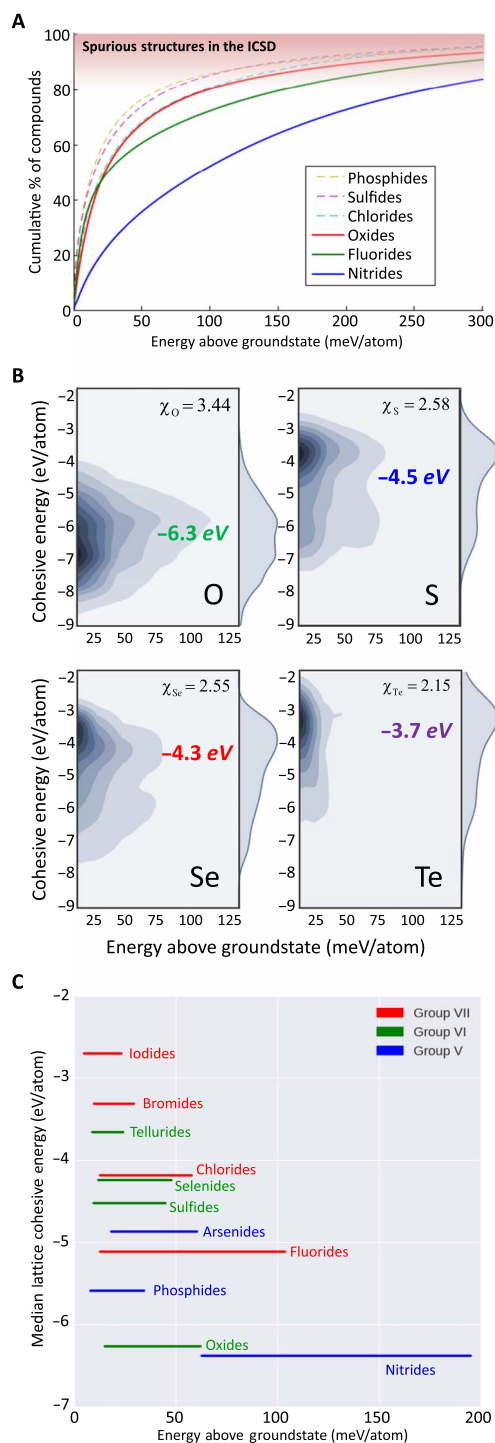


Fig. 1. Influence of chemistry on thermodynamic scale of metastability. (A) Cumulative distribution functions of crystalline metastability for the most-represented chemistries in the Materials Project. Manual investigation reveals that the 20% highest-energy structures in the ICSD do not correspond to observed, crystalline polymorphs. (B) Bivariate sample density maps of metastability versus cohesive energy for group VI compounds. Chemistries with higher electronegativities, χ , exhibit stronger bonds, resulting in greater median cohesive energies and higher accessible crystalline metastabilities. (C) Energy scale of metastability for various chemistries, ordered vertically by the median cohesive energy. Left vertex, median metastability; right vertex, 90th percentile. Within a periodic group, greater lattice cohesivity yields greater crystalline metastability, as strong bonds can lock more metastable crystal structures.

Nitrides are an important class of optoelectronics (35), and the reported synthesizability of highly metastable nitrides from reactive nitrogen precursors (36, 37) suggests that there may be a broad spectrum of promising and technologically relevant metastable nitrides awaiting discovery.

Although our study focuses on the metastability of inorganic crystals, polymorphism and metastability in organic molecular solids is of great technological relevance to pharmaceuticals, organic electronics, and protein folding (7). Our observation relating cohesive energy to metastability could address a deep fundamental question in organic molecular solids: Why do many molecular solids exhibit numerous polymorphs within a small ($\sim 15^\circ\text{C}$) temperature range, whereas inorganic solids often see $>100^\circ\text{C}$ differences between polymorph transition temperatures? The weak intermolecular bonds of molecular solids yield cohesive energies of roughly -100 kJ/mol (12) or -1 eV per molecule, about a third of the cohesion of the weakest class of inorganic solids (iodides; Fig. 2B). This weak lattice cohesion yields a correspondingly small energy scale of organic crystalline metastability (38). When this small energy scale of organic crystalline metastability is coupled with the rich structural diversity arising from high conformational degrees of freedom during molecular packing (39), this inevitably leads to a wide range of accessible polymorphs over a small span of thermodynamic conditions.

Influence of composition

The space of metastable compounds hovers above an energy landscape of equilibrium phases. As chemical elements are added to a thermodynamic system, the complexity of this energy landscape grows. Figure 2A shows an example calculated energy landscape for the ternary Fe-Al-O system, plotted as a convex hull of formation energies referenced to the elemental standard states (see section S1.2 for discussion). We anticipate the thermodynamic metastability of a phase to be different when it is competing against a polymorph—a stable phase of the same composition (Fig. 2A, red stars)—or against a phase-separated state—multiple phases of different compositions (Fig. 2A, purple triangles). In Fig. 2B, we explore this hypothesis by constructing probability distributions of metastability for allotropes, binaries, ternaries, quaternaries, and pentanaries and beyond, grouped by whether the competing equilibrium phase is a polymorph (shaded light) or phase-separated (shaded dark). The relative areas of the shaded regions are proportional to the ratio of entries within each composition.

Figure 2B demonstrates that the more elements present in a metastable compound, the more likely that its competing equilibrium state is phase-separated rather than polymorphic, and that in general, these phase-separating compounds tend to be more metastable than polymorphs. The increased probability for phase separation with increasing number of elements results from a higher likelihood of low-energy decomposition products to exist in a broader chemical space. However, even though this brings about greater thermodynamic risk for phase separation, long-range chemical separation is diffusion-limited, which can be a kinetic barrier that enables the persistence of highly metastable (>70 meV/atom) multinary compounds. Indeed, there are emerging examples that the formation of low-dimensional crystals from multicomponent precursors under diffusion-limited conditions can result in novel crystalline phases that are metastable with respect to phase separation (4, 40–42). In contrast, polymorphic phase transformations occur under constant local composition and thus lack this kinetic barrier of chemical separation, which may rationalize why the energy scale of metastability for

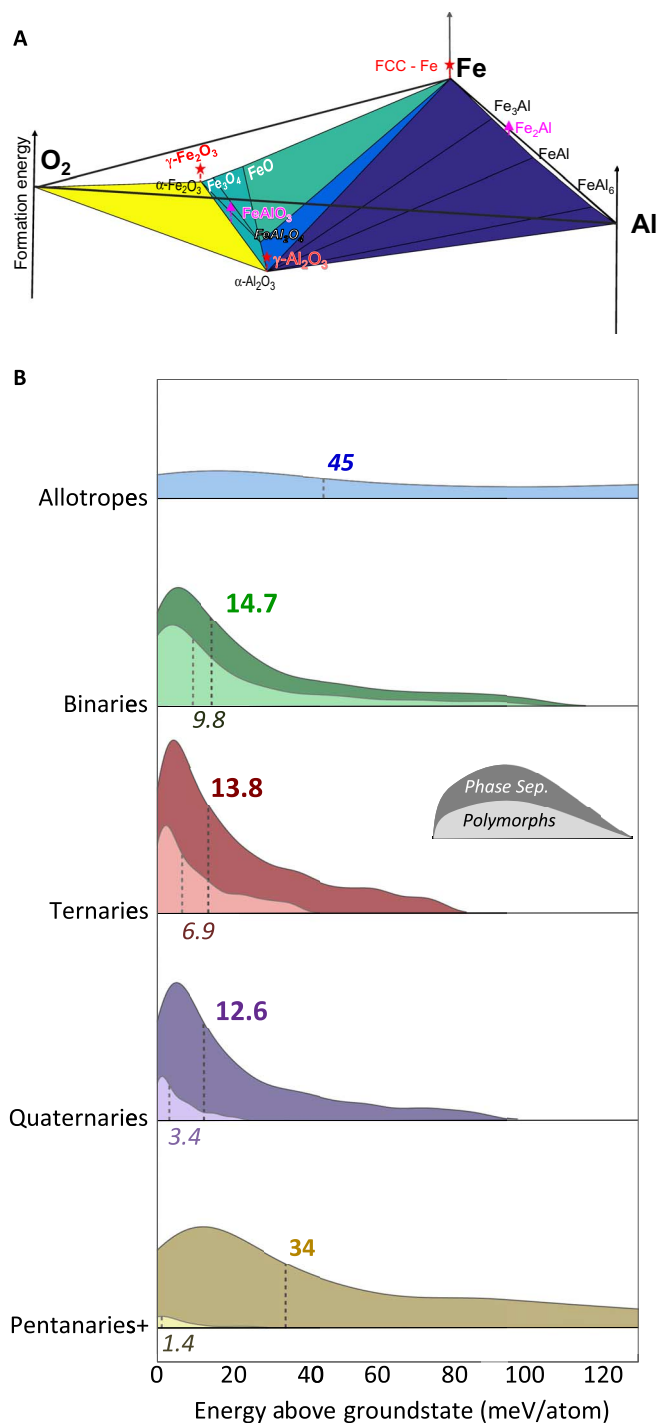


Fig. 2. Influence of composition on the thermodynamic scale of metastability. (A) Computed energy landscape of the ternary Fe-Al-O system with example polymorphic (red stars) and phase-separating (purple triangles) metastable compounds. For example, $\gamma-Fe_2O_3$ is a metastable polymorph, whereas $FeAlO_3$ is metastable with respect to phase separation into $Fe_2O_3 + Al_2O_3$ (energies of metastable phases not drawn to scale). (B) Gaussian kernel density estimates of metastability distributions given the number of different elements in a compound. Dark- and light-shaded regions correspond to phase-separating and polymorphic entries, respectively. Relative areas of the shaded regions correspond to number of entries per category. Top number, median for all entries; bottom number, polymorphs only.

phase-separating compounds is larger in magnitude than that for polymorphic compounds.

We additionally observe in Fig. 2B that for polymorphs, the energy scale of metastability decreases as the number of elements in a compound increases. We conjecture that this is related to the increased scope of structural complexity that accompanies the increase in number of elements in a compound (14), which may yield a denser spectrum of low-energy, thermodynamically-accessible polymorphs. To illustrate this concept: the diversity of possible structures for single-element allotropes is relatively sparse (FCC, BCC, and HCP for metals) with large, discrete energy differences between phases, resulting in a broad and uniform energy distribution of metastable structures. As more elements are incorporated and structural degrees of freedom are made available, more low-energy polymorphs can exist and be formed, resulting in the observed diminishing energy scale of polymorphic metastability. We emphasize that the multinary polymorphs considered here are ordered crystalline phases and not solid solutions, and so their existence cannot simply be explained from configurational entropy.

The competing interactions between polymorphic and phase-separating compounds culminate in ternaries and quaternaries having the lowest observed energy scale of metastability, whereas allotropes (elemental polymorphs), binaries, and pentaries and beyond, all exhibit higher overall metastability. Allotropes and polymorphic binaries are constrained by the few structural degrees of freedom to form low-energy structures, yielding high metastability. Ternaries and quaternaries have both low-energy polymorphic configurations and a limited scope of phase-separated decomposition products, resulting in overall mild metastability. Pentaries and beyond consist primarily of phase-separating metastable structures; thus, their metastabilities are much higher than those of compounds with fewer elements. This final observation suggests an addendum to Pauling's rule of parsimony, which states that "the number of essentially different kinds of constituents in a crystal tends to be small" (43). Although this may be accurate for ground states, this rule may be softened for pentaries and beyond, which can exist in metastable states with numerous unique constituents.

We have thus explicitly quantified the thermodynamic scale of crystalline metastability over the majority of known, computable inorganic solids. We reveal new insights into how the accessible scale of metastability is influenced by specific differences in chemistry, by the number of components in the system, and by what the competing equilibrium state is. Our results provide explicit data-mined evidence to supplement conventional intuition regarding the nature of inorganic crystalline metastability.

DISCUSSION

Metastability and synthesizability

For each composition, there is only one thermodynamically stable ground-state, but there are a potentially infinite number of metastable phases. How does nature decide which of the many possible metastable structures a solid should crystallize in? The formation energies of inorganic solids are on the order of eV/atom (section S3.3), whereas we have found crystalline metastability to be narrowly confined to within tens of meV/atom of the convex hull. It is thus tempting to use energy as a structure-selection mechanism, and to conclude that moderately metastable materials, with an energy above the ground-state commensurate with the observed metastability range ($< \sim 70$ meV/atom, varying by chemistry and

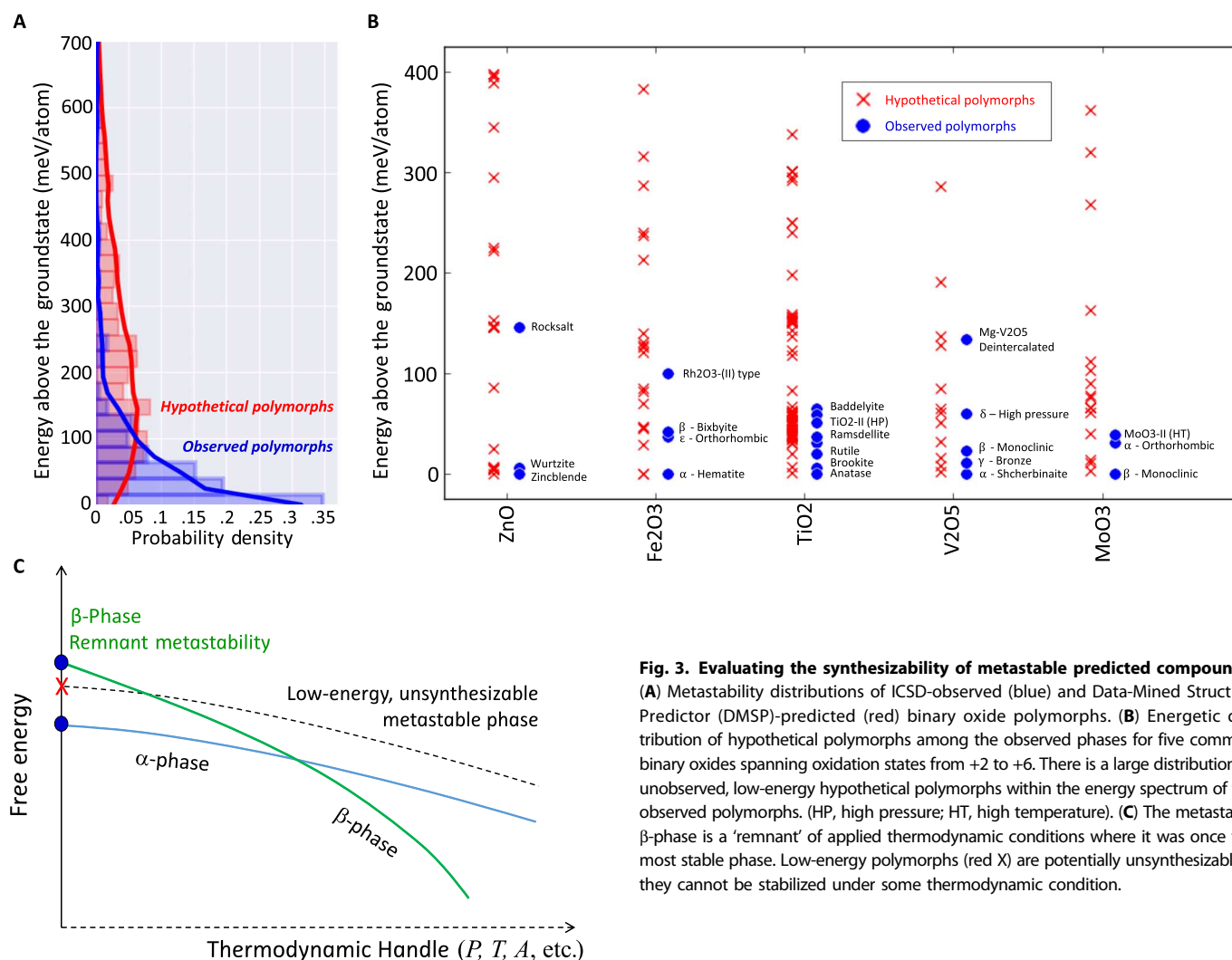


Fig. 3. Evaluating the synthesizability of metastable predicted compounds.

(A) Metastability distributions of ICSD-observed (blue) and Data-Mined Structure Predictor (DMSP)-predicted (red) binary oxide polymorphs. (B) Energetic distribution of hypothetical polymorphs among the observed phases for five common binary oxides spanning oxidation states from +2 to +6. There is a large distribution of unobserved, low-energy hypothetical polymorphs within the energy spectrum of the observed polymorphs. (HP, high pressure; HT, high temperature). (C) The metastable β -phase is a ‘remnant’ of applied thermodynamic conditions where it was once the most stable phase. Low-energy polymorphs (red X) are potentially unsynthesizable if they cannot be stabilized under some thermodynamic condition.

composition), are reasonable candidates for synthesis. Indeed, this assumption is often invoked during the computational discovery and design of new materials (44–46). We will demonstrate that this assumption needs refinement. We find that although observed metastable phases do tend to be low in energy, a low energy of metastability is not a sufficient criterion for synthesis, and that more stringent structure-selection mechanisms must exist.

To better understand which metastable structures can be made, we compare the energy scale of observed metastable phases (ICSD entries) to the energy scale of unobserved metastable phases. As it is, by definition, impossible to probe the energy scale of unobserved metastable phases experimentally, we accomplish this computationally, by generating hypothetical phases from a data-mined structure prediction algorithm (47), which makes rational chemical substitutions based on existing crystal structures to generate unobserved but reasonable novel phases *in silico* (section S1.4). We note that these unobserved phases are based on well-known crystal structures, but in compositions not previously known for those structures. We focus our investigation on binary oxides, which are well-studied due to their geological and technological relevance,

and have been synthesized under a broad variety of conditions in both nature and the laboratory.

Figure 3A shows the metastability distributions between observed ICSD binary oxides and our hypothetical predicted binary oxides. The two distributions are fundamentally different; the distribution for hypothetical polymorphs is broad and Gaussian-like, exhibiting a maximum probability near 150 meV/atom, whereas the distribution for observed polymorphs is negative exponential-like, with the 90th percentile for binary oxides occurring at 94 meV/atom. There is a large spectrum of unobserved, high-energy hypothetical metastable phases, which confirms the intuition that structures with too much excess energy above the ground-state cannot be synthesized. Surprisingly, we also identify many unobserved low-energy hypothetical polymorphs within the energetic spectrum of the observed phases. Figure 3B highlights this observation for five well-studied binary oxides across five different oxidation states: ZnO, TiO₂, Fe₂O₃, V₂O₅, and MoO₃. From lattice phonon calculations (section S1.4), we find that half of the hypothetical structures with a metastability of <100 meV/atom in the Fe₂O₃ system have no imaginary phonon modes, confirming that these low-energy structures are truly metastable, rather than

dynamically unstable (48). There are two possibilities for why these low-energy hypothetical polymorphs are unobserved—either they have not been synthesized yet, or they cannot be synthesized. However, because these particular binary oxides have been so well explored both in nature and in the laboratory, we consider it unlikely that there are *so many* hypothetical polymorphs yet to be synthesized, and rather, it is more likely that they cannot be synthesized. This suggests that a low energy of metastability is not a sufficient condition for synthesizability.

Although energy above the ground-state alone may not be enough to distinguish the synthesizability of a metastable phase, we remember that phases that are stable under applied thermodynamic fields can often be kinetically retained in a metastable state when these thermodynamic fields are removed—for example, diamond carbon or quenched high-temperature austenitic steel. Metastable phases with lower surface energies than the bulk stable phase can also be size-stabilized at the nanoscale (49, 50), resulting in preferential nucleation during crystallization (51, 52). These nano-stabilized phases can persist during crystal growth to sizes much larger than where they are thermodynamically stable (53). We will demonstrate that a mechanism of ‘remnant metastability’—that is, that metastable phases are generally remnants of thermodynamic conditions where they were once the lowest free-energy phase—can account for the magnitude of the energy scale of inorganic crystalline metastability, and possibly why many low-energy hypothetical phases may not be synthesizable.

In general, differences in materials properties between two phases, such as entropy, volume, or surface energy, ΔY , can be exploited by applying a conjugate field, X , to the system (pressure, temperature, or surface area) to change the free-energy difference between two phases, ΔG , according to $d\Delta G = \Delta Y dX$ (Fig. 3C, blue and green lines). Using order of magnitude estimates of ~ 1000 K for temperature, ~ 10 GPa for pressure, and ~ 5000 m²/mol for size (corresponding to ~ 10 -nm particles), an energy difference of 100 meV/atom between polymorphs can be overcome with a ΔS of 10 J/mol·K, a ΔV of 2 Å³/mol, or a $\Delta\gamma$ of 2 J/m². These values are of the same order of magnitude as actually observed materials property differences between thermodynamically competing phases. Other thermodynamic handles, including stress-strain, electric and magnetic fields, and compositional variations, operate in similar energy ranges. It appears that the magnitude of the energy scale of metastability is similar to the strength of thermodynamic handles, suggesting that metastable structure-selection may be thermodynamic in origin, rather than kinetic. Furthermore, if the synthesis of metastable phases requires them to be stable under only some thermodynamic condition, perhaps there are no thermodynamic conditions under which our hypothetical, unobserved phases (12, 16) are ever the lowest free-energy phases (dashed line, Fig. 3C). Finally, the presence of so many unobserved, low-energy phases implies that crystal structure-selection is more than just a sequential energetic cascade through metastable structure-space, as claimed in Ostwald’s original formulation of the “rule of stages” (54), meaning that more sophisticated mechanisms should be invoked when interpreting observations of multistage crystallization (55).

It is compelling that ‘remnant metastability’ can simultaneously constrain the energy scale of crystalline metastability (which is one to two orders of magnitude weaker than the energy scale of formation energies) while also rationalizing why not all low-energy metastable phases can be synthesized. While purely nonequilibrium pathways to metastable phases may certainly still exist, approaching materials synthesis from a thermodynamic perspective of ‘remnant metastability’ can provide more

rational and quantitative strategies for synthesis design than qualitative ‘kinetics’ approaches (10). We thus propose a more refined criterion for the predictive synthesis of novel materials: “Synthesis of novel metastable crystalline phases should target conditions where they are thermodynamically stable, and aim to kinetically retain them to conditions where metastable. If conditions of thermodynamic stability cannot be found, realization of these predicted metastable phases may not be possible.” We emphasize that phase stability can be evaluated under more general thermodynamic potentials than simply under temperature and pressure, and can include other forms of thermodynamic work, such as surface work (size, adsorption), elastic work (epitaxy, stress-strain), electromagnetic work (charge transfer, electrical polarization, magnetic polarization), and chemical work (such as compositional variation) (56)

METHODS

The data in this paper were retrieved from the Materials Project (www.materialsproject.org) (18), using the Python Materials Genomic (pymatgen) package (57) (www.pymatgen.org), through the Materials API framework (32) (www.materialsproject.org/open). DFT calculations for the Materials Project were performed with the Vienna ab initio software package (VASP) (58, 59) using the projector-augmented wave method with the generalized gradient approximation (GGA) within the Perdew-Burke-Ernzerhof (PBE) framework (60). For oxides with strongly correlated electrons, the GGA + U method was used (61), with values of U set by the methodology of Wang *et al.* (24) and refined by Jain *et al.* (23). All compounds were structurally optimized using the FireWorks high-throughput calculations package (62). These DFT calculation parameters were all specified in pymatgen within the MPVaspInputSet settings. For reactions with gas-phase decompositions (O₂, N₂, etc.), the chemical potential of the gas phase was fit using known experimental reaction energies by the method of Wang *et al.* (25). The atomic energy for cohesive energy calculations was attained from ‘atom-in-a-box’ total energy calculations of a single atom in a 10 Å × 10 Å × 10 Å unit cell. More details on the high-throughput calculation methodology can be found in the study by Jain *et al.* (26).

Hypothetical structures were predicted via the data-mined ionic substitution algorithm developed by Hautier *et al.* (47), as implemented in the pymatgen package. Ordered binary compounds from the 2012 version of the ICSD were used as seeding structures for the ionic substitution. Dynamical stability for the six predicted Fe₂O₃ structures within 100 meV/atom of the ground state was evaluated from finite-difference phonon calculations, calculated from the *phonopy* package (63).

Statistical analyses were completed within the NumPy and SciPy packages in Python. Data visualization was performed using the Seaborn statistical data visualization package (64) and the MATLAB Statistics and Machine Learning Toolbox. Further details on the statistical analyses can be found in the Supplementary Materials.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/2/11/e1600225/DC1>

Supplementary Methods

Dataset construction, validation, and provenance

fig. S1. Convex hull in the binary A-B system.

fig. S2. Phonon density of states for lattice-stable low-energy predicted Fe₂O₃ polymorphs.

fig. S3. Phase stability errors should be referenced to adjacent phases in a convex hull, rather than the elemental states.
 fig. S4. Fraction of metastable phases in a convex hull with formation energies jittered by 24 meV/atom random Gaussian variable.
 fig. S5. Probabilities of attaining same stable phases or same stable compositions under a convex hull jittered by random DFT error.
 fig. S6. Cumulative distribution of volume changes in the Materials Project.
 fig. S7. Cumulative distribution of bond length changes in the Materials Project.
 fig. S8. Large change in bond length between the ICSD and computed entry for Ba₂CoO₄.
 fig. S9. Large change in bond length between the ICSD and computed entry for SnF₂.
 fig. S10. Energy distribution of metastable binary oxide polymorphs in the Materials Project, sorted by provenance.
 fig. S11. Gaussian kernel distribution of the formation energies of both stable and metastable ICSD entries in the Materials Project.
 table S1. Data Table for Fig. 1: Metastability by Chemistry
 table S2. Data Table for Fig. 2: Metastability by Composition

REFERENCES AND NOTES

- J. W. Gibbs, On the equilibrium of heterogeneous substances. *Am. J. Sci.* **96**, 441–458 (1878).
- A. Sclafani, J. M. Herrmann, Comparison of the photoelectronic and photocatalytic activities of various anatase and rutile forms of titania in pure liquid organic phases and in aqueous solutions. *J. Phys. Chem.* **100**, 13655–13661 (1996).
- G. P. Nagabhushana, R. Shivaramaiah, A. Navrotsky, Direct calorimetric verification of thermodynamic instability of lead halide hybrid perovskites. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7717–7721 (2016).
- T. F. Kuech, S. E. Babcock, L. Mawst, Growth far from equilibrium: Examples from III-V semiconductors. *Appl. Phys. Rev.* **3**, 040801 (2016).
- M. E. Davis, R. F. Lobo, Zeolite and molecular sieve synthesis. *Chem. Mater.* **4**, 756–768 (1992).
- S. Sanna, V. Esposito, J. W. Andreasen, J. Hjelm, W. Zhang, T. Kasama, S. B. Simonsen, M. Christensen, S. Linderoth, N. Pryds, Enhancement of the chemical stability in confined δ-Bi₂O₃. *Nat. Mater.* **14**, 500–504 (2015).
- J. Bernstein, *Polymorphism in Molecular Crystals*, vol. 14 (Oxford Univ. Press, 2007).
- Z. Li, K. G. Pradeep, Y. Deng, D. Raabe, C. C. Tasan, Metastable high-entropy dual-phase alloys overcome the strength–ductility trade-off. *Nature* **534**, 227–230 (2016).
- J. Gopalakrishnan, Chimie douce approaches to the synthesis of metastable oxide materials. *Chem. Mater.* **7**, 1265–1275 (1995).
- A. Stein, S. W. Keller, T. E. Mallouk, Turning down the heat: Design and mechanism in solid-state synthesis. *Science* **259**, 1558–1564 (1993).
- M. Jansen, A concept for synthesis planning in solid-state chemistry. *Angew. Chem. Int. Ed. Engl.* **41**, 3746–3766 (2002).
- S. L. Price, Why don't we find more polymorphs? *Acta Crystallogr. Sect. B Struct. Sci.* **69**, 313–328 (2013).
- V. Stevanović, Sampling polymorphs of ionic solids using random superlattices. *Phys. Rev. Lett.* **116**, 075503 (2016).
- A. R. Oganov, M. Valle, How to quantify energy landscapes of solids. *J. Chem. Phys.* **130**, 104504 (2009).
- M. Jansen, I. V. Pentin, J. C. Schön, A universal representation of the states of chemical matter including metastable configurations in phase diagrams. *Angew. Chem. Int. Ed. Engl.* **51**, 132–135 (2012).
- M. A. Zwijnenburg, F. Illas, S. T. Bromley, Apparent scarcity of low-density polymorphs of inorganic solids. *Phys. Rev. Lett.* **104**, 175503 (2010).
- S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, O. Levy, The high-throughput highway to computational materials design. *Nat. Mater.* **12**, 191–201 (2013).
- A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
- J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, C. Wolverton, Materials design and discovery with high-throughput density functional theory: The open quantum materials database (OQMD). *JOM* **65**, 1501–1509 (2013).
- S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L. W. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, O. Levy, AFLOWLIB.ORG: A distributed materials properties repository from high-throughput *ab initio* calculations. *Comput. Mater. Sci.* **58**, 227–235 (2012).
- A. Belsky, M. Hellenbrandt, V. L. Karen, P. Luksch, New developments in the Inorganic Crystal Structure Database (ICSD): Accessibility in support of materials research and design. *Acta Crystallogr. Sect. B Struct. Sci.* **58**, 364–369 (2002).
- S. Ping Ong, L. Wang, B. Kang, G. Ceder, Li–Fe–P–O₂ phase diagram from first principles calculations. *Chem. Mater.* **20**, 1798–1807 (2008).
- A. Jain, G. Hautier, S. Ping Ong, C. J. Moore, C. C. Fischer, K. A. Persson, G. Ceder, Formation enthalpies by mixing GGA and GGA+U calculations. *Phys. Rev. B* **84**, 045115 (2011).
- L. Wang, T. Maxisch, G. Ceder, Oxidation energies of transition metal oxides within the GGA+U framework. *Phys. Rev. B* **73**, 195107 (2006).
- L. Wang, T. Maxisch, G. Ceder, A first-principles approach to studying the thermal stability of oxide cathode materials. *Chem. Mater.* **19**, 543–552 (2007).
- A. Jain, G. Hautier, C. J. Moore, S. Ping Ong, C. C. Fischer, T. Mueller, K. A. Persson, G. Ceder, A high-throughput infrastructure for density functional theory calculations. *Comput. Mater. Sci.* **50**, 2295–2310 (2011).
- Y. Wang, S. Curtarolo, C. Jiang, R. Arroyave, T. Wang, G. Ceder, L.-Q. Chen, Z.-K. Liu, *Ab initio* lattice stability in comparison with CALPHAD lattice stability. *Calphad* **28**, 79–90 (2004).
- S. Curtarolo, D. Morgan, G. Ceder, Accuracy of *ab initio* methods in predicting the crystal structures of metals: A review of 80 binary alloys. *Calphad* **29**, 163–211 (2005).
- G. Hautier, S. Ping Ong, A. Jain, C. J. Moore, G. Ceder, Accuracy of density functional theory in predicting formation energies of ternary oxides from binary oxides and its implication on phase stability. *Phys. Rev. B* **85**, 155208 (2012).
- V. Stevanović, S. Lany, X. Zhang, A. Zunger, Correcting density functional theory for accurate predictions of compound enthalpies of formation: Fitted elemental-phase reference energies. *Phys. Rev. B* **85**, 115104 (2012).
- S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, C. Wolverton, The Open Quantum Materials Database (OQMD): Assessing the accuracy of DFT formation energies. *npj Comput. Mater.* **1**, 15010 (2015).
- S. Ping Ong, S. Cholia, A. Jain, M. Brafman, D. Gunter, G. Ceder, K. A. Persson, The Materials Application Programming Interface (API): A simple, flexible and efficient API for materials data based on REpresentational State Transfer (REST) principles. *Comput. Mater. Sci.* **97**, 209–215 (2015).
- R. Niewa, F. J. DiSalvo, Recent developments in nitride chemistry. *Chem. Mater.* **10**, 2733–2752 (1998).
- D. H. Gregory, Structural families in nitride chemistry. *J. Chem. Soc. Dalton Trans.* **3**, 259–270 (1999).
- H. Amano, T. Asahi, I. Akasaki, Stimulated emission near ultraviolet at room temperature from a GaN film grown on sapphire by MOVPE using an AlN buffer layer. *Jpn. J. Appl. Phys.* **29**, L205 (1990).
- D. Fischer, M. Jansen, Synthesis and structure of Na₃N. *Angew. Chem. Int. Ed. Engl.* **41**, 1755–1756 (2002).
- C. M. Caskey, R. M. Richards, D. S. Ginley, A. Zakutayev, Thin film synthesis and properties of copper nitride, a metastable semiconductor. *Mater. Horiz.* **1**, 424–430 (2014).
- J. Nyman, G. M. Day, Static and lattice vibrational energy differences between polymorphs. *CrystEngComm* **17**, 5154–5165 (2015).
- A. J. Cruz-Cabeza, J. Bernstein, Conformational polymorphism. *Chem. Rev.* **114**, 2170–2191 (2013).
- H. Peng, P. F. Ndione, D. S. Ginley, A. Zakutayev, S. Lany, Design of semiconducting tetrahedral Mn_{1-x}Zn_xO alloys and their application to solar water splitting. *Phys. Rev. X* **5**, 021016 (2015).
- R. B. Soriano, I. U. Arachchige, C. D. Malliakas, J. Wu, M. G. Kanatzidis, Nanoscale stabilization of new phases in the PbTe–Sb₂Te₃ system: Pb_mSb_{2n}Te_{m+3n} nanocrystals. *J. Am. Chem. Soc.* **135**, 768–774 (2012).
- A. J. Martinolich, J. A. Kurzman, J. R. Neilson, Circumventing diffusion in kinetically-controlled solid-state metathesis reactions. *J. Am. Chem. Soc.* **138**, 11031–11037 (2016).
- L. Pauling, The principles determining the structure of complex ionic crystals. *J. Am. Chem. Soc.* **51**, 1010–1026 (1929).
- T. F. T. Cerqueira, S. Lin, M. Amsler, S. Goedecker, S. Botti, M. A. L. Marques, Identification of novel Cu, Ag, and Au ternary oxides from global structural prediction. *Chem. Mater.* **27**, 4562–4573 (2015).
- A. Zakutayev, X. Zhang, A. Nagaraja, L. Yu, S. Lany, T. O. Mason, D. S. Ginley, A. Zunger, Theoretical prediction and experimental realization of new stable inorganic materials using the inverse design approach. *J. Am. Chem. Soc.* **135**, 10048–10054 (2013).
- Y. Wu, P. Lazić, G. Hautier, K. Persson, G. Ceder, First principles high throughput screening of oxynitrides for water-splitting photocatalysts. *Energy Environ. Sci.* **6**, 157–168 (2013).
- G. Hautier, C. Fischer, V. Ehrlicher, A. Jain, G. Ceder, Data mined ionic substitutions for the discovery of new compounds. *Inorg. Chem.* **50**, 656–663 (2010).
- V. Ozoliņš, A. Zunger, Theory of systematic absence of NaCl-type (β-Sn-type) high pressure phases in covalent (ionic) semiconductors. *Phys. Rev. Lett.* **82**, 767 (1999).
- A. Navrotsky, Energetics at the nanoscale: Impacts for geochemistry, the environment, and materials. *MRS Bull.* **41**, 139–145 (2016).
- A. S. Barnard, L. A. Curtiss, Prediction of TiO₂ nanoparticle phase and shape transitions controlled by surface chemistry. *Nano Lett.* **5**, 1261–1266 (2005).
- A. Navrotsky, Energetic clues to pathways to biomineralization: Precursors, clusters, and nanoparticles. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12096–12101 (2004).

52. W. Sun, S. Jayaraman, W. Chen, K. A. Persson, G. Ceder, Nucleation of metastable aragonite CaCO_3 in seawater. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3199–3204 (2015).
53. F. Glas, J.-C. Harmand, G. Patriarche, Why does wurtzite form in nanowires of III-V zinc blende semiconductors? *Phys. Rev. Lett.* **99**, 146101 (2007).
54. W. Ostwald, Studien über die Bildung und Umwandlung fester Körper. *Z. Phys. Chem.* **22**, 289–330 (1897).
55. M. H. Nielsen, S. Aloni, J. J. De Yoreo, In situ TEM imaging of CaCO_3 nucleation reveals coexistence of direct and indirect pathways. *Science* **345**, 1158–1162 (2014).
56. R. A. Alberty, Use of Legendre transforms in chemical thermodynamics (IUPAC Technical Report). *Pure Appl. Chem.* **73**, 1349–1380 (2001).
57. S. Ping Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, G. Ceder, Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
58. G. Kresse, J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
59. G. Kresse, J. Furthmüller, Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
60. J. P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
61. V. I. Anisimov, J. Zaanen, O. K. Andersen, Band theory and Mott insulators: Hubbard *U* instead of Stoner *I*. *Phys. Rev. B* **44**, 943–954 (1991).
62. A. Jain, S. Ping Ong, W. Chen, B. Medasani, X. Qu, M. Kocher, M. Brafman, G. Petretto, G.-M. Rignanese, G. Hautier, D. Gunter, K. A. Persson, FireWorks: A dynamic workflow system designed for high-throughput applications. *Concurr. Comput. Pract. E.* **27**, 5037–5059 (2015).
63. A. Togo, I. Tanaka, First principles phonon calculations in materials science. *Scr. Mater.* **108**, 1–5 (2015).
64. M. Waskom, O. Botvinnik, P. Hobson, J. Warmenhoven, J. B. Cole, Y. Halchenko, J. Vanderplas, S. Hoyer, S. Villalba, E. Quintero, A. Miles, T. Augspurger, T. Yarkoni, C. Evans, D. Wehner, L. Rocher, T. Megies, L. P. Coelho, E. Ziegler, T. Hoppe, S. Seabold, S. Pascual, P. Cloud, M. Koskinen, C. Hausler, kjemmett, D. Milajevs, A. Qalieh, D. Allan, K. Meyer, *seaborn*; <http://dx.doi.org/10.5281/zenodo.19108>.

Acknowledgments: W.S. thanks F. Frankel (Massachusetts Institute of Technology) for valuable discussions on data visualization and S. Y. Chan for valuable insights. **Funding:** The data-mining portion of this work was intellectually led by the Materials Project, which was supported by the Department of Energy (DOE) Basic Energy Sciences program under grant no. EDCBEE, DOE contract DE-AC02-05CH11231. The analysis of different forms of metastability was supported by the U.S. DOE, Office of Science, Basic Energy Sciences, under contract no. UGA-0-41029-16/ER392000 as part of the DOE Energy Frontier Research Center “Center for Next Generation of Materials by Design: Incorporating Metastability.” We used computing resources at the Argonne National Laboratory Center for Nanoscale Materials, an Office of Science User Facility, which was supported by the U.S. DOE, Office of Science, Office of Basic Energy Sciences, under contract no. DE-AC02-06CH11357. This research also used computational resources from the Center for Functional Nanomaterials, which is a U.S. DOE Office of Science Facility at Brookhaven National Laboratory, under contract no. DE-SC0012704. **Author contributions:** W.S. and G.C. designed the research and wrote the manuscript. W.S. gathered Materials Project data and performed statistical analyses with A.C.G. Hypothetical structures were generated by S.T.D. and G.H. based on data-mined ionic substitution algorithms. A.J., S.P.O., G.H., K.A.P., and G.C. designed and built the Materials Project. S.P.O., A.J., and G.H. built and calibrated phase stability calculation tools. S.T.D., S.P.O., and W.D.R. wrote the structure-matcher algorithms. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate and reproduce the conclusions in this paper are present in the paper and/or the Supplementary Materials, and are available online for free at www.materialsproject.org and can be accessed with the Materials API at www.materialsproject.org/open, using the pymatgen package (www.pymatgen.org). Additional data related to this paper may be requested from the authors.

Submitted 4 November 2015

Accepted 20 October 2016

Published 18 November 2016

10.1126/sciadv.1600225

Citation: W. Sun, S. T. Dacek, S. P. Ong, G. Hautier, A. Jain, W. D. Richards, A. C. Gamst, K. A. Persson, G. Ceder, The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2**, e1600225 (2016).

This article is published under a Creative Commons license. The specific license under which this article is published is noted on the first page.

For articles published under [CC BY](#) licenses, you may freely distribute, adapt, or reuse the article, including for commercial purposes, provided you give proper attribution.

For articles published under [CC BY-NC](#) licenses, you may distribute, adapt, or reuse the article for non-commercial purposes. Commercial use requires prior permission from the American Association for the Advancement of Science (AAAS). You may request permission by clicking [here](#).

The following resources related to this article are available online at <http://advances.sciencemag.org>. (This information is current as of November 19, 2016):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
<http://advances.sciencemag.org/content/2/11/e1600225.full>

Supporting Online Material can be found at:
<http://advances.sciencemag.org/content/suppl/2016/11/14/2.11.e1600225.DC1>

This article **cites 62 articles**, 5 of which you can access for free at:
<http://advances.sciencemag.org/content/2/11/e1600225#BIBL>

Science Advances (ISSN 2375-2548) publishes new articles weekly. The journal is published by the American Association for the Advancement of Science (AAAS), 1200 New York Avenue NW, Washington, DC 20005. Copyright is held by the Authors unless stated otherwise. AAAS is the exclusive licensee. The title *Science Advances* is a registered trademark of AAAS