Rational solid-state synthesis routes for inorganic materials

Muratahan Aykol,* Joseph H. Montoya, and Jens Hummelshøj

Toyota Research Institute, Los Altos, CA 94022, United States

E-mail: murat.aykol@tri.global

Abstract

Rational solid-state synthesis of inorganic compounds is formulated as catalytic nucleation on crystalline reactants, where contributions of reaction and interfacial energies to the nucleation barriers are approximated from high-throughput thermochemical data, and structural and interfacial features of crystals, respectively. Favorable synthesis reactions are then identified by a Pareto analysis of relative nucleation barriers and phase-selectivities of reactions leading to the target. We demonstrate the application of this approach in reaction planning for solid-state synthesis of a range of compounds, including the widely-studied oxides $LiCoO_2$, $BaTiO_3$ and $YBa_2Cu_3O_7$, as well as other metal oxide, oxyfluoride, phosphate and nitride targets. Pathways for enabling retrosynthesis of inorganics are also discussed.

Introduction

Solid-state synthesis of inorganic materials lacks a general theory to facilitate rational planning and selection of reactions.[1–4] Formation of a new solid phase often occurs through nucleation and growth, phenomena that are driven by a complex interplay of bulk, surface and interface thermodynamics and transport,[5,6] hindering any straightforward, step-by-step

reconstruction of crystals from smaller components as in organic retrosynthesis.[7] Hence, machine-learning and text-mining based approaches are being sought towards enabling predictive synthetic capability for inorganic solids,[8–14] accompanied by first-principles studies providing in-depth analyses of reaction mechanisms in individual systems or higher-level synthesizability trends.[15–18] Lack of a generally applicable rational synthesis planning framework is considered the missing link for realization of computer-designed functional inorganic materials.[1,4,19]

While a reaction involving two or more solids may eventually become rate-limited by diffusion of reacting species towards the reaction zone as the product thickens,[20–22] the emergent product phase at the *onset* of the reaction is often controlled by nucleation. This concept of *phase selection through nucleation* has the potential to enable design of rational synthesis routes for inorganics. In fact, choosing starting materials that are not only favorable for the reaction thermodynamics, but can also provide a surface for heterogeneous, catalytic nucleation of the target phase is a familiar concept in solid-state chemistry.[21,23] However, a quantitative treatment of nucleation with *ab-initio* or atomistic computations is challenging, and as such has not yet yielded a practical method for synthesis route prediction for inorganic solids. Here we show that high-throughput thermochemical data and crystal structure features of materials can be combined to approximate the relative favorability of solid-state syntheses within the well-known classical nucleation theory (CNT).[6,23–26] The resulting approach for rational planning of inorganic solid-state synthesis routes (hereafter referred as PIRO) can rapidly yield a set of plausible precursors and subsequent solid-state reactions for a target inorganic compound, and is broadly applicable.

Theory

Starting with the steady-state CNT description,[6,23–26] for a given reaction k with solid reactants $\alpha_i \in \{\alpha_1, \alpha_2, ..., \alpha_n\}$, we write the rate of heterogeneous nucleation of a target phase β

on the surface of a reactant α_i at temperature T as:

$$J_{\alpha_i \to \beta} = J_0 \cdot \exp[-\Delta G^*_{\alpha_i \to \beta}/k_B T] \tag{1}$$

where J_0 is the pre-exponential factor, and $\Delta G^*_{\alpha_i \to \beta}$, is the critical nucleation barrier, which can be written for spherically-shaped nuclei as:

$$\Delta G^*_{\alpha_i \to \beta} = (16\pi/3)\gamma^3_{\beta v}[\Delta G_k]^{-2} f(S_{\alpha_i \to \beta}) \tag{2}$$

Here $\gamma_{\beta v}$ is the surface energy of phase β, ΔG_k is the free energy of reaction k per volume of the cluster of phase β being formed, and the factor $f(S_{\alpha_i \to \beta})$ varies from 0 to 1 and quantifies the reduction in the barrier from the homogenous (uncatalyzed) limit on the surface of α_i as a function of $S_{\alpha_i \to \beta}$ (the cosine of the contact angle $\theta_{\beta\alpha_i}$ between substrate α_i and β) described by:

$$S_{\alpha_i \to \beta} = \cos(\theta_{\beta\alpha_i}) = (\gamma_{\alpha_i v} - \gamma_{\beta\alpha_i})/\gamma_{\beta v}. \tag{3}$$

γ values correspond to interfacial energies between the subscript phases (or surface energy in case of v). For β nucleating as a spherical cap, f can be written as a monotonic function: $(2 - 3S_{\alpha_i \to \beta} + S^3_{\alpha_i \to \beta})/4$.

While ΔG_k is relatively straightforward to estimate from thermochemical data, a general, quantitative assessment of Eq. 2 is hindered by the extreme difficulty in computation or experimental measurement of interfacial and surface energies in Eq. 3. However, as $\gamma_{\beta v}$ is constant for a given target β, a direct approximation of $S_{\alpha_i \to \beta}$ itself, instead of the interfacial energy terms, would be sufficient to obtain *relative* values of nucleation barriers for different reactions k using Eq. 2.

To identify the reactions that may have a relatively higher catalytic nucleation effect for synthesis of the target phase, we need to look for those with small $f(S_{\alpha_i \to \beta})$. This condition can simply be achieved in CNT by maximizing $S_{\alpha_i \to \beta}$ through maximizing the surface energy

of α and minimizing $\beta-\alpha$ interface energy, which however would require quantitative assessments of these two terms. There is nevertheless a special case of Eq. 3 within the physical boundaries of catalytic CNT that we can target; i.e. $S_{\alpha_i\to\beta} \to 1$. This condition is tractable through the *energy ratios* in Eq. 3, and is in fact a *ubiquitous* special case that is in line with the well-known guidelines in solid-state reactions[21,27] that choosing starting materials (i) structurally and chemically (or topochemically) similar and (ii) epitaxially commensurate to a target compound can be beneficial for a reaction to succeed. In view of Eq. 3, we see that argument (i) would help obtain similar magnitudes for surface energies (and hence $\gamma_{\alpha_i v}/\gamma_{\beta v}$ remain close to ≈ 1; see Fig. S1), and (i) and (ii) together would facilitate a small interfacial energy, $\gamma_{\beta\alpha_i}$ (and hence $\gamma_{\beta\alpha_i}/\gamma_{\beta v} \to 0$ at the ideal limit).

In light of these empirical arguments, since our goal is to simply separate the favorable reactions from others, we assume a simple tangent plane approximation near $S_{\alpha_i\to\beta} = 1$ to capture the deviation from this ideal upper bound as:

$$S_{\alpha_i\to\beta} \approx 1 - q^{sim}_{\alpha_i,\beta} - q^{epi}_{\alpha_i,\beta} \tag{4}$$

where $q^{sim}_{\alpha_i,\beta}$ and $q^{epi}_{\alpha_i,\beta}$ are distance metrics normalized to [0,1] that quantify similarity and epitaxial matching of α_i and β, respectively, and are easily obtainable from the crystal structures, as explained later. This approximation becomes exact at the limit $\alpha_i = \beta$, as long as $q \to 0$ for both cases. Therefore, with reasonable approximations to the similarity and epitaxy related metrics, Eq. 4 should return values closer to 1 for α_i that has high catalytic potency for nucleation of β and should otherwise be smaller when such potency is low or uncertain within the boundaries of the framework. In effect, CNT-based catalytic nucleation formulation for $\Delta G^*_{\alpha_i\to\beta}$ is reduced to a practical filter that prioritizes catalytically and thermodynamically favorable synthesis reactions over others, using only data easily accessible from thermochemical databases and crystal structures. The pre-exponential factor J_0 depends on the phases involved in catalytic nucleation,[25] but entering as an exponential

term, $\Delta G^*_{\alpha_i \to \beta}$ would be the predominant factor in Eq. 1 influencing the relative nucleation rates towards β, and hence suitable for qualitatively filtering reactions on the basis of their nucleation potency. Still, one specific contribution to J_0 that we should consider is the rate factor capturing the kinetics of cluster growth with an exponential (Arrhenius-type) temperature dependence on an activation barrier $\Delta E_{\alpha_i \to \beta}$. Considering that solid-state synthesis almost exclusively takes place at moderately high-temperatures and $\Delta E_{\alpha_i \to \beta}$ is difficult to quantify for solid-state components, one option could be to neglect this term and test the use of nucleation barrier directly. Diffusion can often be facilitated by high-temperature (e.g. Tammann's rule[28]) and use of starting materials that decompose (e.g. carbonates) upon heating. As another option, $\Delta E_{\alpha_i \to \beta}$ can be used as a parameter deduced from potential diffusion bottlenecks, and lumped together with $\Delta G^*_{\alpha_i \to \beta}$ to obtain an effective barrier to rank reactions, $\Delta G^{**}_{\alpha_i \to \beta} \approx \Delta G^*_{\alpha_i \to \beta} + \Delta E_{\alpha_i \to \beta}$. For example, presence of a topotactical relationship between α_i and β may lead to easier atomic arrangements at the reaction interface,[21] which may be crudely captured as $\Delta E_{\alpha_i \to \beta} \propto q^{sim}_{\alpha_i,\beta}$. Ultimately, for a given candidate reaction k, we assign the minimum catalytic barrier to nucleation we find among reactants, $\Delta G^{**}_k = \min\{\Delta G^{**}_{\alpha_1 \to \beta}, \Delta G^{**}_{\alpha_2 \to \beta}, ..., \Delta G^{**}_{\alpha_n \to \beta}\}$.

The relative potency of reactions for catalytic nucleation of a given target β can now be compared through their ΔG^{**}_k values. However, an important question remains open: what is the selectivity of reactants of a given reaction for forming the target β vs. other *parasitic* phases? A high-fidelity assessment of selectivity through comparison of nucleation rates of different products would require a *quantitative* description of surface/interface energy terms entering Eqs. 2 and 3, and hence is intractable. Therefore, we turn to a simpler data-driven and thermodynamically motivated heuristic that can be evaluated easily to capture the extent of phase competition: the larger the number of other products the reactants of a reaction k can be rebalanced to yield with a favorable free energy, the higher the chances of that reaction producing parasitic phases. We prefer reactions that minimize this number as they likely have higher selectivity for target β. As illustrated in Fig. 1, reactants further

away from the target on composition and energy axes can potentially lead to formation of a larger number of viable phases. For example, reactants of R0, which has the largest reaction energy for forming the target, can also yield three other compounds in the same space, whereas R5 sees no competition but at the expense of a smaller reaction energy. We can see that minimization of this selectivity metric competes with ΔG_k, and in effect, the nucleation barrier.

This particular choice for a "selectivity" metric is motivated by several factors. First, while not sufficient; a favorable reaction energy ($\Delta G_k < 0$) is a necessary condition for formation. If likelihood of formation (p) is mostly uniformly distributed among such $N_{comp}(k)$ competing phases accessible from reactant set k, the expected number of parasitic phases $N_{comp}(k) \times p$ is $\propto N_{comp}(k)$, which can hence serve as a reasonable measure of selectivity. In fact, Sun et al.[15] showed that experimentally-known inorganic materials span wide energy ranges above the *convex-hull* (i.e. ground states and/or their mixtures), with broad, heavy-tailed probability distributions. In other words, p is not expected to dramatically vary among materials thermodynamically accessible (target or competing) from the relatively stable starting materials we consider in this work. Second, the exact ratio of reactants do not have much bearing on the microscopic nucleation process,[21] and hence all phases within the compositional envelope enclosed by reactants should be considered in competition. Lastly, this metric relies on data availability, and should be thought as a *relative* metric to compare reactions and not as an absolute handle on the actual number of parasitic phases that may form. The overall aim is to efficiently distinguish more selective reactions from less selective ones (on which we later provide empirical evidence and further analyses).

Hence, we now have a data-driven framework (outlined in Fig. 1) that reasonably captures the underlying physics for the onset of phase transformations leading to synthesis of a target phase from CNT, and in turn factors in both free energies of reactions and potential catalytic effects of their reactants for targeted synthesis of a phase β, as well as likelihood of reactions to produce parasitic phases. Finding the most favorable reactions for β is reduced to a Pareto

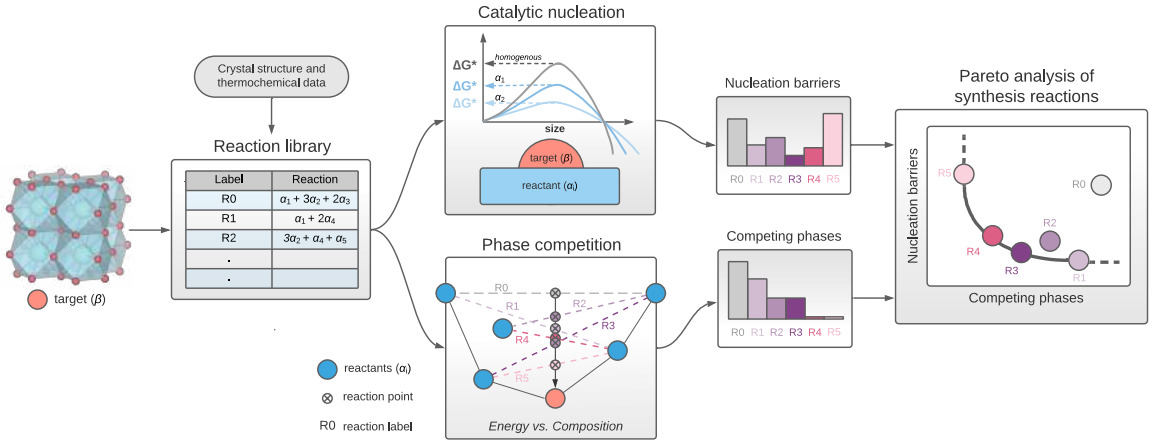

Figure 1: Schematic illustration of the main computational steps in rational solid-state synthesis planning approach, PIRO. Given a target phase β, one begins by building a reaction library enumerating all possible reactions leading to the target phase, then analyzing each of these reactions on the basis of their relative nucleation barriers (Catalytic nucleation) and the number of competing phases along the reaction pathway (Phase competition). With a Pareto analysis, reactions having optimal or nearly-optimal trade-offs between the two metrics can be determined.

optimality problem of minimizing the respective nucleation barrier and phase competition metrics. Using this guidance, a prospective set of reagents can be chosen such that it has the highest predicted rate of formation at a given number of potential side reactions, or the highest ideal selectivity for a given relative rate.

Methods

Formation enthalpy and crystal structure information for all solid phases are obtained from the Materials Project (MP).[29] In calculating the reaction free energies, temperature effects are considered primarily through the enthalpy and entropy changes for the gaseous reactants and products (e.g. O_2, CO, CO_2, H_2O, N_2 etc.) using the experimental standard values obtained from Barin[30] and NBS thermochemical tables.[31] Change in gas free energy with partial pressure with respect to the standard pressure of 1 atm is considered by adding $RT\ln(P)$, where R is the gas constant, to the gas chemical potentials. We neglect the

finite temperature enthalpy and entropy effects or PV contributions for solid phases, as they are often negligible compared to that of the gas phases. Since carbonates are particularly relevant for the examples in the following sections, a correction factor was fitted and added per CO_3^{2-} (or equivalently per CO_2, see Fig. S2) which ensures that experimental energies of decomposition reactions of metal carbonates to their respective oxides are reproduced accurately with DFT energy values obtained from the MP database in analogy with prior reference corrections.[32,33]

Epitaxial-matching is computed as the minimal coincident area between a pair of phases using the method by Zur and McGill,[34] for planar interfaces with miller indices up to 2, using the implementation by Ding et al.[35] in pymatgen,[36] normalized by 10^3 Å^2. Similarity (topochemical) between a pair of phases is measured as the quantile of the Euclidean distance between 273-dimensional Voronoi-tessellation and composition-statistics based standardized feature vectors of Ward et al.[37] using the MatMiner[38] package. Both metrics are desired to be small for nucleation purposes, and their normalizations ensure consistency with their intended use as q in Eq.4. In coupling barrier and diffusion terms in ${\Delta G_{\alpha_i \to \beta}}^{**}$, we adopt $\Delta E_{\alpha_i \to \beta} \approx C q^{sim}_{\alpha_i,\beta}$, treating C as a hyper-parameter set at 10 eV and assuming $\gamma_{\beta v} \approx 2$ J/m^2, which in effect serve the purpose of practically weighting structural and chemical similarity more when approximating the barrier, with effective weights of each term in ${\Delta G_{\alpha_i \to \beta}}^{**}$ adjusted through proper selection of these parameters near typical values, rather than treating them as absolute terms.

For a given target phase, we find all possible balanced reactions that would lead to it from available entries in the same chemical system (validating that the composition matrices of reactions have proper effective ranks). Addition of an extra element to the chemical space covering the target compound is allowed, so as to include common decomposable reactants/precursors (e.g. metal carbonates, hydroxides, nitrates etc.) or enable combustion-type reactions. We do not explicitly test potential destabilization/decomposition of a reactant phase (e.g. carbonate) under the prescribed temperature-pressure conditions (which

would be straightforward to do), in order not to prematurely exclude potentially useful reactions and assume such information would be factored in upon selection of reaction conditions by the chemist. While the framework is not restricted to it, in the current examples, we consider reactions that yield one solid phase (target) phase. A second phase is allowed to evolve, if it is gaseous (e.g. CO_2), and reactions are balanced accordingly. In that sense, current implementation primarily pertains to high-temperature, ceramic, combustion or other solid-state synthesis/mechanical-attrition routes, including atmosphere controlled scenarios, that can be treated within these reaction constructions. The competing phase reactions are generated with the same constraints described above with single solid products, as most reactions with multiple solid products would be separable to those of single targets. Any *inseparable* multiple-solid product reactions would require co-nucleation of two or more solids via complex, poorly-understood "barrier layer" arrangements and transport requirements that make them less likely beyond certain templates,[21] to the best of our knowledge. The competing phases are searched in the chemical space of the target when respective data in MP is abundant, otherwise extended to one degree smaller sub-spaces. We confirm later that the current procedure for calculating this selectivity metric is adequate for our purposes and in agreement with experiments.

The interactive versions of the recommendation plots reported in this work can be found in Supporting Information, which provide easier navigation of the reaction spaces than static plots and tables. The python library supporting the current framework will be open-sourced at http://github.com/TRI-AMDD/ (upon publication). This software library provides an easy-to-use interactive tool to generate similar plots for other systems.

Results and discussion

Overview of the practical use of PIRO approach

In the following sections, we present several case studies using the PIRO approach, starting with three well-known functional metal oxides: the ferroelectric $BaTiO_3$, the common Li-ion battery cathode $LiCoO_2$, and superconducting cuprate $YBa_2Cu_3O_7$. The vast literature on synthesis of these technological materials allows us the perform an in-depth validation of the presented synthesis route recommendation approach. Next, we provide numerous case studies in less common, and chemically diverse set of systems to demonstrate general applicability of the approach.

For every synthesis planning application, we need to specify which materials should be considered as potential starting materials. For example, we find there to be 340 different materials (considering also the carbonated forms) in the MP database at the time of writing that can be used to generate about 1.4×10^5 balanced reactions for synthesis of $BaTiO_3$. Majority of these reactions are likely to be impractical simply based on what is available to the chemist or at least easily-accessible, and hence it is prudent to introduce several filters to help choose a viable subset of potential reactants. To start with, we may check whether a particular material is (i) stable or within a certain distance of the energy-composition convex-hull (e.g. 10 meV/atom) and/or (ii) experimentally-sourced (for which we use the existence of an MP entry in the Inorganic Crystal Structure Database, ICSD[39] as a proxy). Additionally, we may filter reactant candidates by (iii) whether they are *intermediates* (e.g. ternary compounds other than the target tetragonal $BaTiO_3$ in Ba-Ti-O chemical space) and (iv) whether a particular compound has a less common oxidation state (e.g. peroxides or superoxides). When there are too many reactants, we may focus on those that were reported as precursors in literature-mined synthesis recipes,[10] which yields a more restrictive set than (ii). Overall, as a general recipe for single-step reaction identification, we start with a set of *standard* reactants comprised of experimentally-known, fully stable, simple (i.e. no

intermediates or unusual oxidation states) materials in the parent chemistry of the target, and incrementally add more materials to the library as needed, such as any essential metastable materials (e.g. rutile TiO_2, which is metastable in DFT[40]) or peroxides (e.g. BaO_2). For the retrosynthetic analysis, we follow a more specific strategy as explained later. Lastly, no such filters are needed for the phase competition metric as it aims for quantifying as broadly as possible the relative numbers of thermodynamically accessible phases from a set of reactants.

Case study: Ferroelectric $BaTiO_3$

The tetragonal perovskite $BaTiO_3$ phase, BTO (MP: 5986) is a widely-studied industrial-grade ferroelectric material. High-temperature reactions (generally between 800°C and 1300°C) with common precursors like $BaCO_3$ or BaO, and TiO_2 are used to conventionally make BTO.[41–46] The reactions pertaining to these exact reactants are found to be highly favorable by PIRO (Fig.2(a) and (b), and Table 1) when one considers standard reactants (i.e. elements, simple binary compounds and carbonates of Ba and Ti). Alternative combustion synthesis routes using the peroxide as Ba source (BaO_2) have also been reported,[47–49] mainly mixing metallic Ti and TiO_2 and utilizing the highly oxidizing BaO_2 to propagate the reaction. Such a peroxide route is recovered near the lower bound of Pareto frontier as R9 in Fig.2(c). We also find alternative peroxide driven reactions more favorable in terms of phase competition (such as R5, R6, R7 or R8).

Conventional routes that use $BaCO_3$ and TiO_2 as starting materials often yield impurity phases such as Ba_2TiO_4 and $BaTi_2O_5$, but these intermediates ultimately convert to BTO during synthesis.[44–46,50–54] This common observation is consistent with (i) the notable catalytic effect of BTO on precipitation of these intermediates from the same starting materials (Figs. S3 and S4, in line with phase evolutions in Refs.[45,51–53]), and (ii) the movement of the Pareto front in Fig.2(d) towards more favorable regions in reactions that use such intermediate compounds as reactants (e.g. R11). While these routes would not be practical for BTO as they require *a priori* synthesis of complex intermediates (often more difficult to

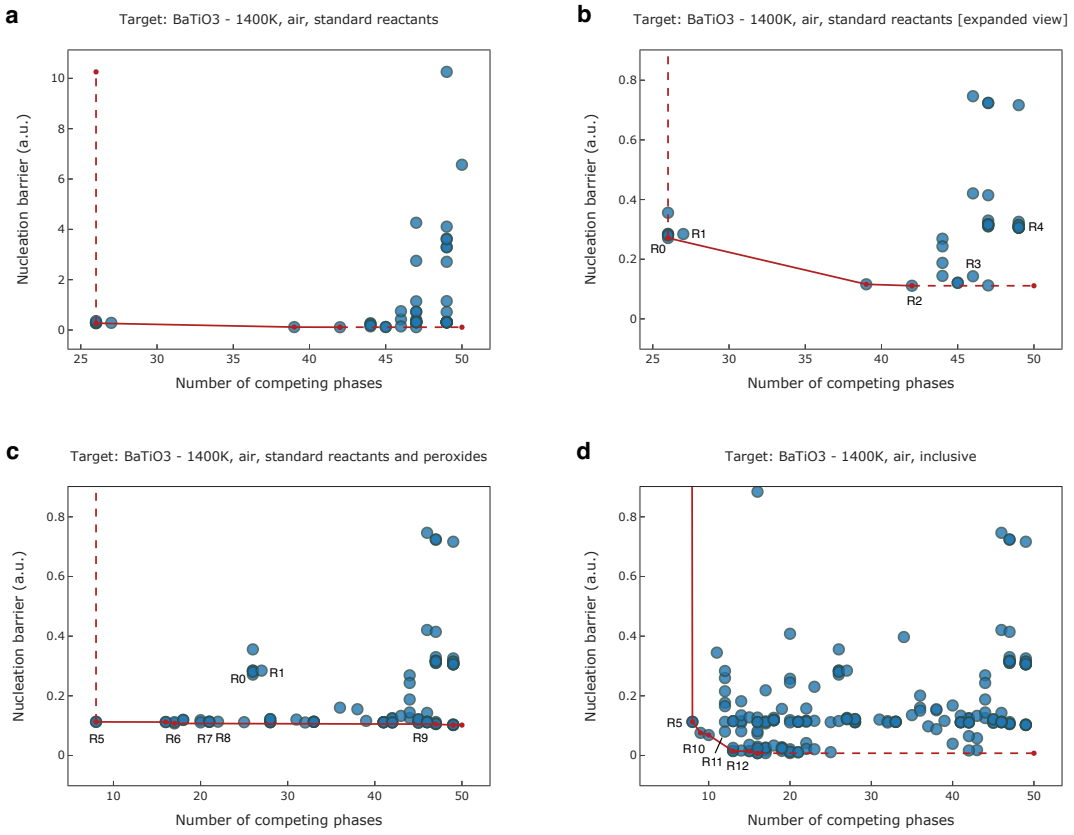

Figure 2: Route planning plots for synthesis of $BaTiO_3$ under typical thermodynamic conditions using various sets of starting materials. Standard reactants include elements, stable binary compounds and carbonates. The set labeled as "inclusive" contains all stable compounds in the Ba-Ti-O-C system. All starting materials are restricted to the experiment-sourced structures present in the ICSD. Pareto fronts in each panel are shown as lines. The recommended procedures are those relatively closer to the origin and points forming or near the Pareto frontier. Plots zoom into regions of interest for clarity, and hence certain high barrier reactions may not be shown.

synthesize than BTO[55] in solid-state), they indicate that deliberately controlling participation of intermediates through favorable reactions can be useful as a synthesis strategy. We will later formalize generation of such multi-step pathways through a retrosynthetic analysis approach.

Table 1: Reactions highlighted in Fig.2 for solid-state $BaTiO_3$ synthesis. Minus (-) sign implies gas release alongside product. The particular form of a compound can be inferred from its MP entry number provided after its formula. Reactions are not necessarily ordered or displayed based on favorability.

Label	**Reaction**
R0*	1.0 BaO(1342) + 1.0 TiO_2(2657)
R1*	1.0 TiO_2(2657) + 1.0 $BaCO_3$(4559) + -1.0 CO_2
R2	0.5 Ti_2O_3(458) + 0.25 O2 + 1.0 BaO(1342)
R3	0.5 Ti_2O_3(458) + 0.25 O2 + -1.0 CO_2 + 1.0 $BaCO_3$(5504)
R4	1.0 TiC(631) + 2.0 O2 + -2.0 CO_2 + 1.0 $BaCO_3$(4559)
R5	1.0 BaO_2(1105) + 1.0 TiO(1203)
R6	0.5 BaO_2(1105) + 0.5 Ti_2O_3(458) + 0.5 BaO(1342)
R7	1.0 BaO_2(1105) + 0.3333 Ti_2O_3(458) + 0.3333 Ti(72)
R8	1.0 BaO_2(1105) + 0.5 Ti_2O_3(458) + -0.25 CO_2 + 0.25 C(569304)
R9*	1.0 BaO_2(1105) + 0.5 Ti(72) + 0.5 TiO_2(2657)
R10	0.3333 Ti_2O_3(458) + 0.3333 BaO_2(1105) + 0.3333 Ba_2TiO_4(3397)
R11	0.5 BaO(1342) + 0.5 $BaTi_2O_5$(3943)
R12	0.5 $BaTi_2O_5$(558159) + 0.25 Ba(122) + 0.25 BaO_2(1105)

*Established synthesis route (or a derivative of such a route) recovered by the planning system.

Case study: Layered $LiCoO_2$

High-temperature (layered) form of $LiCoO_2$, or LCO (MP: 22526) is a widely used cathode material for Li-ion batteries, which can be synthesized in solid-state using Li_2CO_3 as Li source, and Co, CoO, Co_3O_4 or $CoCO_3$ as Co source.[56–61] We first generate recommendation plots for LCO synthesis using *standard* reactants as input (Fig.3(a), and Table 2). We find that such conventional reactions reported in the literature are immediately recovered as favorable (in terms of Pareto optimality) among reactions devised from this precursor library. In particular, the dominance of reactions R0 and R1 is in agreement with the reported mechanism that most Co-sources first lead to formation of Co_3O_4 (or CoO, depending on

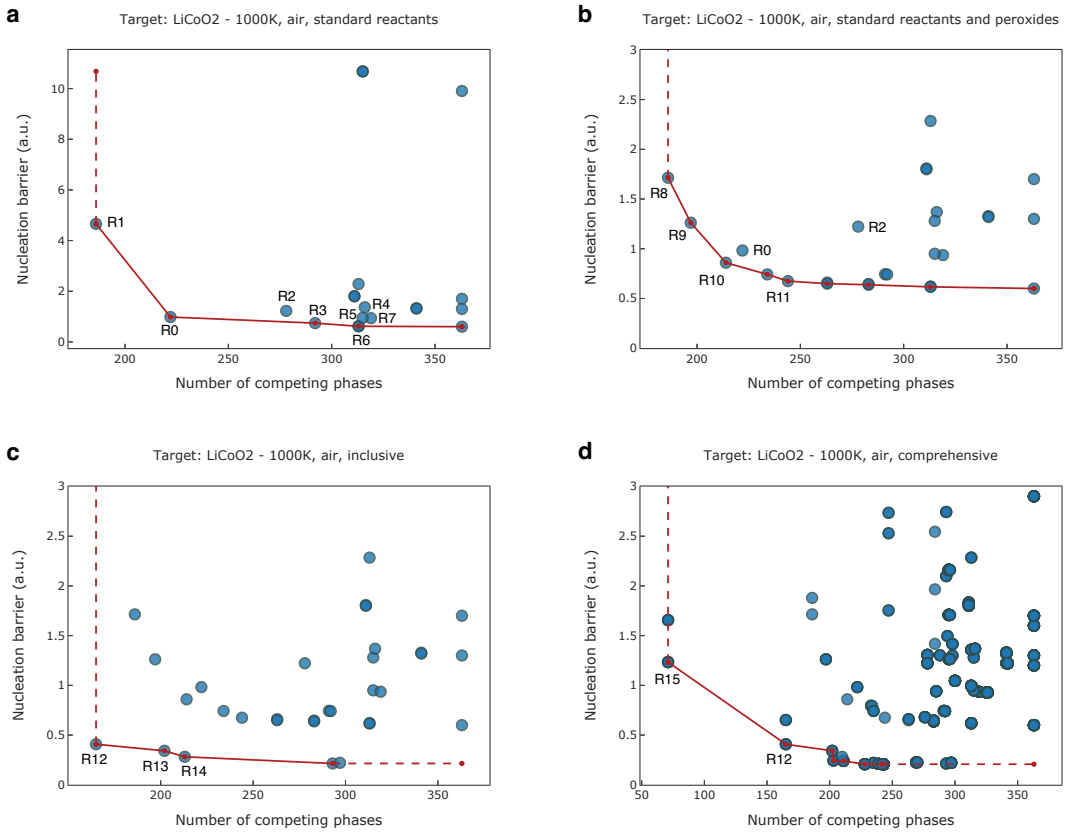

Figure 3: Route planning plots for synthesis of high-temperature (HT) form of layered $LiCoO_2$ under typical thermodynamic conditions using various sets of starting materials. Standard reactants include elements, stable binary compounds and carbonates. The set labeled as "inclusive" contains all stable compounds in the Li-Co-O-C system, and the set labeled as "comprehensive" further adds metastable compounds that are within 10 meV/atom of the convex-hull. All starting materials are restricted to the experiment-sourced structures present in the ICSD. Pareto fronts in each panel are shown as lines. The recommended procedures are those relatively closer to the origin and points forming or near the Pareto frontier. Plots zoom into regions of interest for clarity, and hence certain high barrier reactions may not be shown.

the temperature) which then reacts with the available Li precursor.[57,59,60] Next, we test whether *less conventional* but potentially more favorable routes exist by allowing peroxide reactants (Fig.3(b)) and find that several Li_2O_2 driven reactions occupy the space between the favorable/conventional R0 and R1. One of the most favorable routes in Fig.3(b); a peroxide route based on Li_2O_2 and CoO (R8) was in fact reported by Johnston et al.[62] as a successful option for LCO synthesis, validating the predictions.

In the remaining panels of Fig.3, we show how the landscape of candidate synthesis reactions for LCO changes as we allow intermediates as reactants. For instance, the low profile Pareto frontier that appears in Fig.3(c) is dominated by antifluorite Li_6CoO_4 bearing reactions (Table 2). The favorability of such reactions (e.g. R12, R13 and R14) in Fig.3(c) is straightforward to reason: Li_6CoO_4 has a topotactic relationship with LCO and reduces phase competition due to compositional proximity, and conversion between these two solid phases is in fact readily observed in experiments.[63] We find that synthesis of Li_6CoO_4 from standard reactions stands out as highly favorable even when we include peroxides (See Fig. S5), which results primarily from the catalytic effect of being iso-structural with Li_2O, and is consistent with experiments.[64] Hence, we see this two-step route that emerges from PIRO predictions is in fact viable as each of its steps are corroborated by prior experimental results. Overall, while discovery of multi-step synthesis routes may not be critical for LCO, such routes can be designed through PIRO to achieve a more controlled pathway to the target, which provides further motivation to formulate a retrosynthetic planning, as explained later.

In Fig.3(d), we observe that further addition of metastable phases to the precursor library does not alter the position of the majority of the Pareto front, and above reactions remain among the most favorable. Only at the lower phase competition end does the nominal cathode reaction for LCO batteries (R15: CoO_2 + Li) become part of the Pareto front. This is not a practical synthesis reaction *per se* (as the practical route to CoO_2 itself is delithiation of LCO[65] in an electrochemical setup), but signifies the plausibility of the structural and chemical relationship measures used in PIRO, which clearly exists between CoO_2 and

LCO. Interestingly, certain reactions involving Co_3O_4 remain slightly more favorable on the nucleation axis, mainly because it provides the same degree of epitaxial similarity through the closed-packed oxygen sub-lattice, with the improved chemical similarity through pre-existing Co^{3+}, which may also explain the general success of reported Co_3O_4 routes in LCO synthesis.[57,59,60]

Table 2: Reactions highlighted in Fig.3 for solid-state HT-$LiCoO_2$ synthesis. Minus (-) sign implies gas release alongside product. The particular form of a compound can be inferred from its MP entry number provided after its formula. Reactions are not necessarily ordered or displayed based on favorability.

Label	**Reaction**
R0*	0.5 Li_2O(1960) + 0.0833 O_2 + 0.3333 Co_3O_4(18748)
R1*	1.0 CoO(22408) + 0.5 Li_2O(1960) + 0.25 O_2
R2*	1.0 CoO(22408) + -0.5 CO_2 + 0.25 O_2 + 0.5 Li_2CO_3(3054)
R3*	0.0833 O_2 + 0.3333 Co_3O_4(18748) + -0.5 CO_2 + 0.5 Li_2CO_3(3054)
R4*	0.5 Li_2O(1960) + -1.0 CO_2 + 0.25 O_2 + 1.0 $CoCO_3$(21434)
R5*	-0.5 CO_2 + 1.0 Co(54) + 0.75 O_2 + 0.5 Li_2CO_3(3054)
R6*	1.0 Li(135) + 0.3333 O_2 + 0.3333 Co_3O_4(18748)
R7*	-1.5 CO_2 + 0.25 O_2 + 0.5 Li_2CO_3(3054) + 1.0 $CoCO_3$(21434)
R8*	1.0 CoO(22408) + 0.5 Li_2O_2(841)
R9**	0.5 Li_2O_2(841) + -1.0 CO_2 + 1.0 $CoCO_3$(21434)
R10**	0.3333 Li_2O(1960) + 0.1667 Li_2O_2(841) + 0.3333 Co_3O_4(18748)
R11**	0.5 Li_2O_2(841) + 0.25 Co_3O_4(18748) + 0.25 Co(54)
R12	0.25 O_2 + 0.8333 CoO(22408) + 0.1667 Li_6CoO_4(18925)
R13	0.1111 O_2 + 0.1667 Li_6CoO_4(18925) + 0.2778 Co_3O_4(18748)
R14	0.2 Li_2O_2(841) + 0.1 Li_6CoO_4(18925) + 0.3 Co_3O_4(18748)
R15***	1.0 CoO_2(32686) + 1.0 Li(135)

*Established synthesis route (or a derivative of such a route) recovered by the planning system. **Promising route suggested by the planning system. ***Standard overall reaction of LCO containing battery with a nominal Li anode.

Case study: Superconducting $YBa_2Cu_3O_7$

The chemistry that covers $YBa_2Cu_3O_7$, or YBCO (MP: 20674) has been thoroughly studied in the literature due to the high-temperature superconducting properties of this class of cuprates. The common conventional routes for YBCO synthesis involve Y_2O_3, CuO (or Cu_2O) and $BaCO_3$ (or BaO),[66–68] and are found to be Pareto optimal (R0, R1 and R3) in

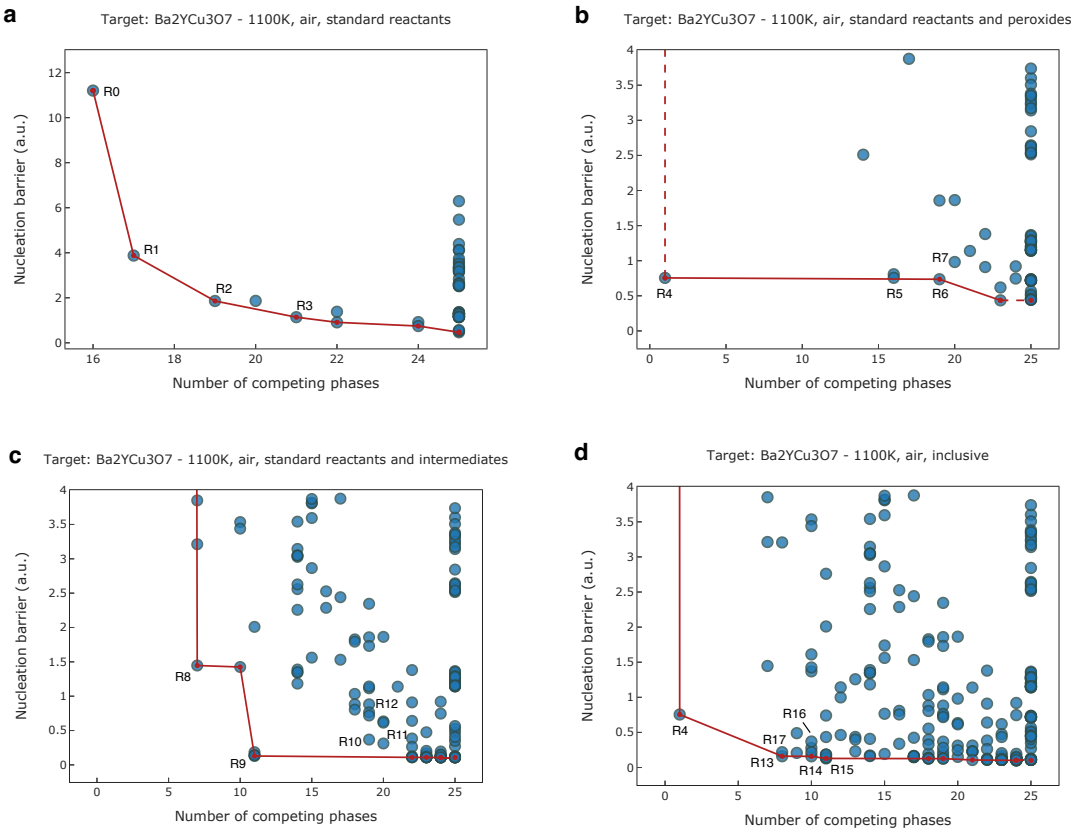

Figure 4: Route planning plots for synthesis of tetragonal $YBa_2Cu_3O_7$ under typical thermodynamic conditions using various sets of starting materials. Standard reactants include elements, stable binary compounds (excluding Y-Cu and Y-Ba alloys for simplicity) and carbonates. The set labeled as "inclusive" expands the set of standard reactants and peroxides to include ternary intermediates (e.g. $Y_2Cu_2O_5$). All starting materials are restricted to the experiment-sourced structures present in the ICSD. Pareto fronts in each panel are shown as lines. The recommended procedures are those relatively closer to the origin and points forming or near the Pareto frontier. Plots zoom into regions of interest for clarity, and hence certain high barrier reactions may not be shown.

Table 3: Reactions highlighted in Fig.4 for solid-state YBCO synthesis. Minus (-) sign implies gas release alongside product. The particular form of a compound can be inferred from its MP entry number provided after its formula. Reactions are not necessarily ordered or displayed based on favorability.

Label	**Reaction**
R0*	3.0 CuO(1692) + 0.25 O_2 + 0.5 Y_2O_3(2652) + 2.0 BaO(1342)
R1*	1.0 O_2 + 1.5 Cu_2O(361) + 0.5 Y_2O_3(2652) + 2.0 BaO(1342)
R2*	3.0 Cu(30) + 1.75 O_2 + 0.5 Y_2O_3(2652) + 2.0 BaO(1342)
R3*	-2.0 CO_2 + 3.0 CuO(1692) + 0.25 O_2 + 0.5 Y_2O_3(2652) + 2.0 $BaCO_3$(4559)
R4*	2.0 BaO_2(1105) + 1.5 Cu_2O(361) + 0.5 Y_2O_3(2652)
R5**	3.0 CuO(1692) + 0.5 Y_2O_3(2652) + 0.5 BaO_2(1105) + -1.5 CO_2 + 1.5 $BaCO_3$(4559)
R6**	1.5 Cu(30) + 1.5 CuO(1692) + 0.5 Y_2O_3(2652) + 2.0 BaO_2(1105)
R7*	2.0 BaO_2(1105) + 3.0 CuO(1692) + -0.75 O_2 + 0.5 Y_2O_3(2652)
R8	-0.5 CO_2 + 0.6667 $Ba_2(CuO_2)_3$(615789) + 0.1667 Ba(122) + 0.5 $BaCO_3$(4559) + 0.5 $Y_2Cu_2O_5$(
R9	2.25 O_2 + -1.75 CO_2 + 1.0 $Ba_2(CuO_2)_3$(615789) + 0.25 Y_4C_7(1200885)
R10*	0.25 O_2 + 2.0 CuO(1692) + -2.0 CO_2 + 2.0 $BaCO_3$(4559) + 0.5 $Y_2Cu_2O_5$(2882)
R11	1.25 O_2 + -2.0 CO_2 + 2.0 Cu(30) + 2.0 $BaCO_3$(4559) + 0.5 $Y_2Cu_2O_5$(2882)
R12	2.0 Cu(30) + 1.25 O_2 + 2.0 BaO(1342) + 0.5 $Y_2Cu_2O_5$(2882)
R13	1.0 Cu(30) + 2.0 BaO_2(1105) + 0.5 Cu_2O(361) + 0.5 $Y_2Cu_2O_5$(2882)
R14	1.5 Cu(30) + 2.0 BaO_2(1105) + 0.5 CuO(1692) + 0.5 $Y_2Cu_2O_5$(2882)
R15	1.75 BaO_2(1105) + 0.25 BaY_2O_4(3952) + 1.25 Cu_2O(361) + 0.25 $Y_2Cu_2O_5$(2882)
R16	1.0 BaO_2(1105) + 0.5 Cu_2O(361) + 0.5 $Y_2Cu_2O_5$(2882) + 1.0 $BaCuO_2$(997034)
R17	1.5 BaO_2(1105) + 0.5 Cu_2O(361) + 0.5 $Ba(CuO)_2$(7374) + 0.5 $Y_2Cu_2O_5$(2882)

*Established synthesis route (or a derivative of such a route) recovered by the planning system.

PIRO recommendations for the use of standard (i.e. elements, simple binary compounds and carbonates of metals, excluding metal-alloys) precursors (Fig.4(a) and Table 3), none being any better than the other simultaneously on both nucleation and selectivity axes. A relatively higher barrier for a particular route such as R0 compared to other routes does not preclude it from nucleation (as every reaction shown already has favorable energy; and hence is thermodynamically feasible), it simply implies that nucleation is likely to be *relatively* more difficult. In fact, achieving phase purity and completion through these conventional routes is known to be challenging (requiring long calcination times, intermittent grindings, oxygenation etc.), and therefore many alternative routes have been pursued for solid-state synthesis of YBCO.[69,70] Among these, peroxide (BaO_2) driven combustion routes with different Cu sources, such as metallic Cu,[71] Cu_2O[72,73] and CuO[70,74–76] have received notable attention as they enable shorter, often single-step synthesis. In agreement with these reports, peroxide-driven routes from PIRO (Fig.4(b) and Table 3) dominate the Pareto-optimal regions, predicted to outperform the common precursor routes above. In particular, the BaO_2-Cu_2O-Y_2O_3 route (R4), which was reported to be highly favorable for producing phase-pure YBCO with no carbonate impurities,[72,73] is predicted as one of the most-favorable routes in Fig. 4(b). The method overall remains predictive even when a reaction's in−situ pathway may potentially involve intermediate steps, which can explained by the explicit *localized* consideration of each possible reaction front between the reactants and target for catalytic nucleation, isolated from any other concurrent precipitations. In Fig. 4(c), we observe that addition of ternary intermediates to standard reactant sets introduces many new reactions. These intermediates appear frequently during synthesis of YBCO,[68,77–80] and can participate in reactions that lead to YBCO formation (e.g. R10 was reported previously[68]). In Fig. 4(d), we see that the previously confirmed, simple peroxide route R4 remains one of the most viable options even when more complex reactions are designed using both peroxides and intermediates (also Table 3). We will revisit YBCO in a later section when designing a strategy for retrosynthesis of inorganics using PIRO.

Table 4: Assessment and validation of relatively complex solid-state synthesis routes using PIRO. Only the solid reactants are listed, and any balancing input or output gases are not shown for simplicity. Stability denotes the distance to the convex-hull in MP database at the time of the study in units of meV/atom.

Target	Stability	Reactants	Exp. Ref.	PIRO display
Ba_2YTaO_6	0	$YTaO_4$, $BaCO_3$	[81]	Fig. 5
$CaSnSiO_5$	8	$CaSnO_3$, SiO_2	[82]	Fig. S6
$LiNa_5Mo_9O_{30}$	0	Li_2MoO_4, MoO_3, Na_2CO_3	[83]	Fig. S6
Sr_2NiWO_6	0	$NiWO_4$, $SrCO_3$	[84]	Fig. 5
$RuSr_2GdCu_2O_8$	21	Sr_2GdRuO_6, Cu_2O	[85]	Fig. S6
$BaV_{10}O_{15}$	0	$Ba_3V_2O_8$, VO, V_2O_3	[86]	Fig. 5
Sr_2CoMoO_6	0	$CoMoO_4$, $SrCO_3$	[87]	Fig. S6
$K_3Y(VO_4)_2$	0	KVO_3, KO_2, Y_2O_3	[88]	Fig. S6
$Na_2Te_3Mo_3O_{16}$	13	Na_2TeO_3, MoO_3, TeO_2	[89]	Fig. S6
$K_2Mo_9S_{11}$	21	K_2MoS_4, MoS_2, Mo	[90]	Fig. S7
$SrNbO_3$	8	$Sr_5Nb_4O_{15}$, Nb	[91,92]	Fig. 5
$RbFe(MoO_4)_2$	0	$Fe_2(MoO_4)_3$, Rb_2MoO_4	[93]	Fig. S7
$K_2V_3P_4O_{17}$	64	$K_4V_2O_7$, P_2O_5, VO_2, V_2O_3	[94]	Fig. 5
$Pr_3BSi_2O_{10}$	0	Pr_6O_{11}, B_2O_3, SiO_2, Si_3N_4	[95]	Fig. S7

Beyond standard systems: Case studies on complex routes and diverse chemistries

Having done an in-depth validation of the PIRO with the abundant literature data available on BTO, LCO and YBCO, we now turn to testing the approach on *less-common* targets, particularly focusing on examples where the reactions complexities are high, targets are metastable or difficult to form, and crucially, *negative* examples of synthesis attempts can be drawn. Our main strategy to find such stringent test cases starts by searching the literature-mined synthesis recipes[10] and extracting reports where the reactant set includes a *non-standard* ingredient, which we define as a ternary or higher-order compound (excluding trivial cases like carbonate precursors). We further augment this set of examples with a manual search over the solid-state literature that is deliberately kept at a cursory level for broader testing, the only criteria being the reported compound to have a clear solid-state route and a matching entry in MP. We list the main results of this analysis in Table 4, with corresponding planning plots shown in Fig. 5 and Figs. S6 and S7. Using these examples, we

can validate PIRO under complex synthesis scenarios, and derive insights from its predictions in light of experimental data.

The first distinct group of examples that emerges from Table 4 involves successful syntheses of quaternary or higher order compounds from ternary precursors, and for which syntheses with more *standard* routes were also reported. Zhou et al. showed that cubic Ba_2YTaO_6 can made from a $YTaO_4$ precursor as well as the binary oxides/carbonates of each metal.[81] Similarly, Mouyane et al. prepared $CaSnSiO_5$ using $CaSnO_3$,[82] which is also realizable through SnO_2, $CaCO_3$ and SiO_2.[96] $LiNa_5Mo_9O_{30}$ is accessible not only through a Li_2MoO_4 route, but also through alkali carbonates and MoO_3.[83,97] In all these examples, we consistently find both types of routes to be on or near the Pareto fronts in PIRO, with similar nucleation barriers and ternary-precursor routes having slightly higher selectivities. These studies do not have sufficient information to compare the selectivity itself, but that will be possible in the next group.

The second distinct group involves cases with both *successful* and *negative* attempts for the synthesis of a same target, an exceedingly difficult but critical validation set to extract from literature. In our context, the *negative* examples constitute either reports on an entirely unsuccessful attempt, or as is more common, attempts that are measurably inferior in terms of phase purity in comparison to other known routes. For solid-state synthesis of Sr_2NiWO_6, Blum et al.[84] reported that the conventional NiO, WO_3 and $SrCO_3$ routes commonly yield secondary phases $SrWO_4$ and Sr_2WO_5, which can be mitigated with the $NiWO_4$-$SrCO_3$ route. Similarly, a high percentage of $SrMoO_4$ impurity is unavoidable when preparing Sr_2CoMoO_6 from $SrCO_3$, MoO_3 and Co_3O_4,[98] whereas the $CoMoO_4$ precursor route of Lufaso et al.[87] does not yield any impurities detectable in XRD. Bauernfeind et al.[99] reported formation of $RuSr_2GdCu_2O_8$ from $SrCO_3$ and simple oxides, but stressed that the resulting samples were not single phase. More recently, Artini et al.[85] demonstrated that precursors Sr_2GdRuO_6 and Cu_2O attained from simpler reagents lead to a single phase $RuSr_2GdCu_2O_8$ with no detectable impurities. With a series of systematic attempts, Ridgley and Ward[100] were able

to induce mixtures containing Sr_xNbO_3 phases using $SrCO_3$ (or SrO) as Sr source, and some combination of Nb_2O_5 and Nb metal, but a phase-pure $SrNbO_3$ was not attainable. In comparison, recent studies showed monophasic $SrNbO_3$ can be synthesized starting with a $Sr_5Nb_4O_{15}$ precursor and Nb metal.[91,92] Brauchle and Huppertz[95] reported a novel route for synthesis of $Pr_3BSi_2O_{10}$ that involved Si_3N_4 and SiO_2 concurrently, and mentioned alternative routes with different combinations of precursors were unsuccessful, leading to formation of other Pr compounds. For all examples cited above, we find clear agreement with PIRO predictions, where the *negative* examples are consistently inferior to the *successful* routes. That is, PIRO helps identify the complex routes that, in comparison to simpler routes, deliver higher selectivity by forming a narrower compositional envelope around the target while maintaining favorable nucleation tendencies.

There are several examples that deserve a closer inspection of the nucleation-selectivity trade-off. For synthesis of $BaV_{10}O_{15}$ from Ba-carbonate, under reducing H_2 conditions, V_2O_3 was found to be a better precursor than V_2O_5 as the latter did not lead to phase pure $BaV_{10}O_{15}$,[101] which the authors attributed to increased water release. But further corroborating the experimental observations, PIRO indicates the V_2O_5 route also suffers from very poor selectivity compared to V_2O_3. In a follow up study, authors noted that the $BaCO_3$-V_2O_3-H_2 route still posed purity issues, and introduced an entirely new route that uses $Ba_3V_2O_8$, VO and V_2O_3.[86] Both types of routes are on the Pareto front in PIRO; the new route displaying a *very high* selectivity, at the expense of nucleation potency. We found other cases showcasing a similar compromise, e.g. $RbFe(MoO_4)_2$ synthesis from ternary Rb and Fe molybdates,[93] or $K_2Mo_9S_{11}$ from a path that involves K_2MoS_4.[90] The success of these types of routes hinges on the fact when there is very low risk for parasitic phase formation with the highly narrow compositional envelope, and a lower but still favorable driving force (high selectivity and of Pareto front), researchers are still able to promote nucleation using external stimuli; e.g. by employing temperatures higher than typically possible through standard precursors.[86,90,93,95]

Above we discussed the most informative examples that emerged from our analyses. There are many cases that show a simple agreement between PIRO and a reported experimental route, including synthesis of $Na_2Te_3Mo_3O_{16}$ from a Na_2TeO_3 precursor and binary oxides,[89] $K_3Y(VO_4)_2$ from KVO_3, peroxide KO_2 and Y_2O_3,[88] $K_2V_3P_4O_{17}$ from $K_4V_2O_7$ and binary oxides.[94] Many *conventional* examples were also studied, where routes with standard reactants are in highly favorable locations; e.g. Ca_2CrSbO_6 by Retuerto et al.[102] (Fig. S7), Sr_2FeO_3F by Galasso and Darby[103] (Fig. S7), $NaTi_8O_{13}$ by Akimoto and Takei[104] (Fig. S7) and Ca_3VN_3 through a reaction of binary nitrides by Vennos and DiSalvo[105] (Fig. 5), further showcasing broad applicability of the framework.

Solid-state retrosynthetic pathways

In case studies targeting BTO, LCO and YBCO, we concluded that the PIRO approach can help design multi-step synthesis pathways. Similarly, capturing of routes in Table 4 that involve (intermediate) reactants that themselves can be made from simpler precursors support the viability of designing multi-step pathways. In this section, we seek to formulate a more formal strategy for finding viable reaction trees that can exploit intermediate structures to realize the solid-state synthesis of a target inorganic compound from commonly used precursors, akin to the retrosynthetic planning of organic molecules.[7,106] And in analogy with organic synthesis, we need to define a set of rules and codify the chemists' heuristics to arrive at tractable synthesis plans. We will use YBCO as an example throughout this section.

As an essential requirement to be able to plan multi-step routes with PIRO, reaction steps arriving at any intermediates, precursors or the target itself need to be executed in sequence and independently. Similar to our general strategy, we limit the space of starting materials considered for reactions to ICSD-based MP entries that are stable, and explicitly add any experimentally-known metastable phase only when needed (e.g. a particular precursor or a known intermediate that is metastable in MP). Even then, inclusion of complex

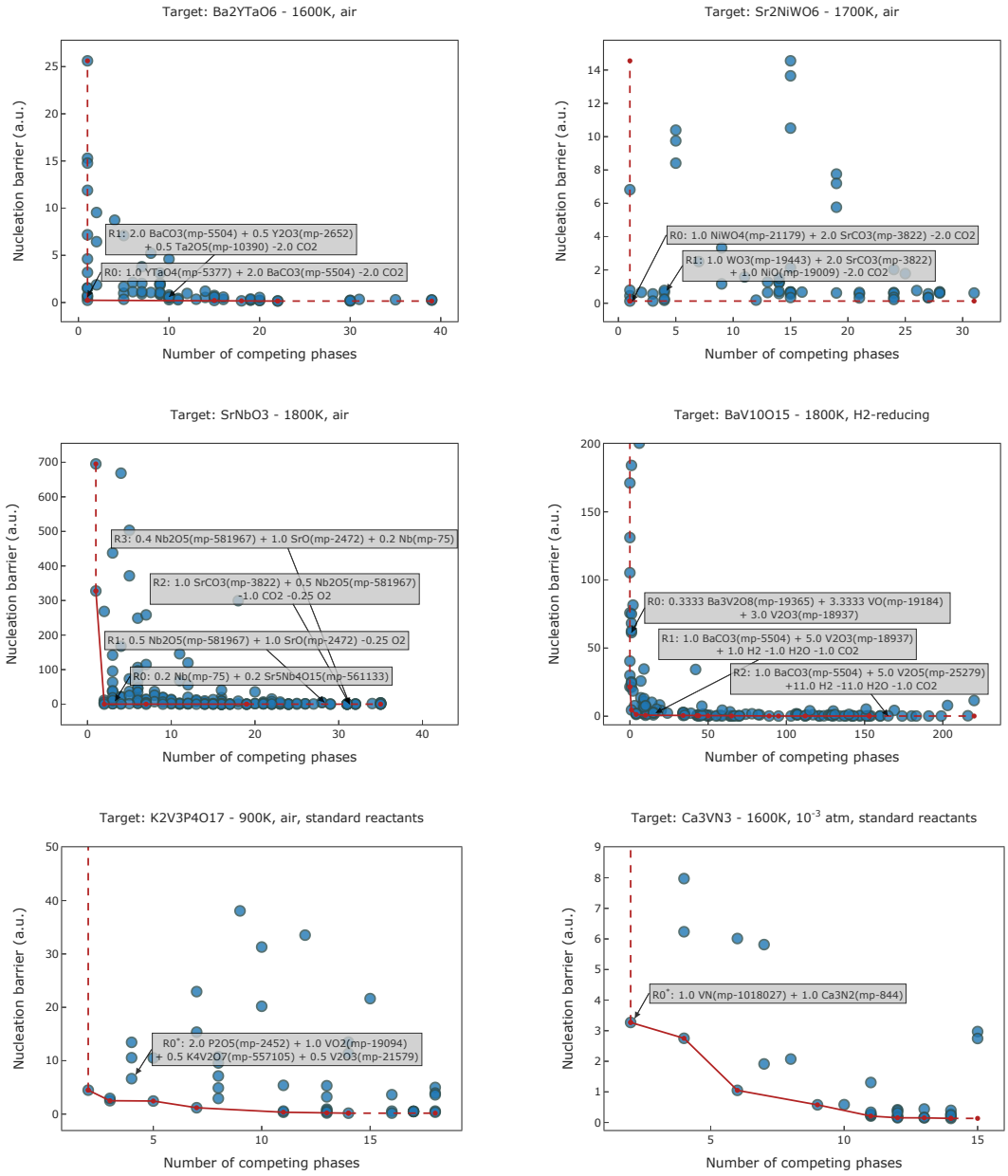

Figure 5: Route planning plots for solid-state synthesis various inorganic compounds selected from the solid-state chemistry literature as described in the text.

Table 5: Synthesis reactions obtained from retrosynthetic analysis of YBCO. Reaction labels list the MP id of the target compound and a reaction number, separated by a colon (e.g. 20674:R0).

Label	**Reaction**
YBCO	
20674:R0	0.5 Y_2O_3(2652) + 2.0 BaO_2(1105) + 1.5 Cu_2O(361)
20674:R1	0.5 O_2 + 0.5 Y_2O_3(2652) + 0.5 Cu_2O(361) + 2.0 $BaCuO_2$(997034)
20674:R2	1.0 Cu(30) + 0.75 O_2 + 0.5 Y_2O_3(2652) + 2.0 $BaCuO_2$(997034)
20674:R3	0.25 O_2 + 2.0 CuO(1692) + 0.5 Y_2O_3(2652) + 1.0 Ba_2CuO_3(8790)
20674:R4	0.75 O_2 + 0.5 Y_2O_3(2652) + 1.0 Cu_2O(361) + 1.0 Ba_2CuO_3(8790)
20674:R5	0.5 BaO(1342) + 1.5 $Ba(CuO)_2$(7374) + 1.0 O_2 + 0.5 Y_2O_3(2652)
20674:R6	1.5 $Ba(CuO)_2$(7374) + 0.75 O_2 + 0.5 Y_2O_3(2652) + 0.5 BaO_2(1105)
20674:R7	2.0 BaO(1342) + 0.25 O_2 + 2.0 CuO(1692) + 0.5 $Y_2Cu_2O_5$(2882)
20674:R8	1.5 $Ba(CuO)_2$(7374) + 1.0 O_2 + 0.5 Y_2O_3(2652) + -0.5 CO_2 + 0.5 $BaCO_3$(4559)
20674:R9	2.0 Cu(30) + 1.25 O_2 + 0.5 Y_2O_3(2652) + 1.0 Ba_2CuO_3(8790)
Intermediates	
$Ba(CuO)_2$	
7374:R0	-1.0 CO_2 + 1.0 Cu_2O(361) + 1.0 $BaCO_3$(4559)
7374:R1	1.0 $BaCO_3$(5504) + -1.0 CO_2 + 1.0 Cu_2O(361)
7374:R2	2.0 CuO(1692) + 1.0 Ba(122)
7374:R3	1.0 BaO_2(1105) + 2.0 Cu(30)
7374:R4	1.0 Cu_2O(361) + 1.0 BaO(1342)
$BaCuO_2$	
997034:R0	1.0 BaO_2(1105) + 1.0 Cu(30)
997034:R1	1.0 CuO(1692) + -1.0 CO_2 + 1.0 $BaCO_3$(4559)
997034:R2	1.0 CuO(1692) + -1.0 CO_2 + 1.0 $BaCO_3$(5504)
997034:R3	1.0 BaO_2(1105) + 1.0 CuO(1692) + -0.5 O_2
997034:R4	1.0 BaO_2(1105) + 0.5 Cu_2O(361) + -0.25 O_2
Ba_2CuO_3	
8790:R0	1.0 CuO(1692) + -2.0 CO_2 + 2.0 $BaCO_3$(4559)
8790:R1	1.0 CuO(1692) + -2.0 CO_2 + 2.0 $BaCO_3$(5504)
8790:R2	1.0 CuO(1692) + -1.0 O_2 + 2.0 BaO_2(1105)
8790:R3	-0.5 O_2 + 2.0 BaO_2(1105) + 1.0 Cu(30)
8790:R4	0.5 Cu_2O(361) + -0.75 O_2 + 2.0 BaO_2(1105)
$Y_2Cu_2O_5$	
2882:R0	1.0 Y_2O_3(2652) + 1.0 O_2 + 2.0 Cu(30)

reactants such as peroxides and intermediates can easily yield hundreds of feasible reactions to consider in the analysis, as seen in Fig.4(d). Besides, a set of reactions that have too many intermediate phases to make would lead to reaction trees that are too wide or deep, and we would want to keep auxiliary syntheses of any intermediate phase that is not available as a commercial reagent to a minimum. In addition, the higher the number of solid reactants in a reaction, the higher the likelihood of solid-phase transport to become a bottleneck in progress of the reaction. In light of these arguments, we introduce the following heuristics for carrying out the retrosynthetic analysis for inorganics using PIRO:

1. We choose reactions leading to the target to have at most one intermediate compound that has not been listed as a *known precursor* in the literature among the reactants, which we assess using the literature-parsed solid-state precursor dataset of Kononova et al.[10] The remainder of the selected reactants should either be part of the Kononova dataset and/or are elements or binary compounds (or e.g. their carbonated versions). We exclude reactions that require making metal-alloys (e.g. YCu_2), except metal-carbides or elemental graphite to attain carbothermal[107] and/or self-sustaining high-temperature synthesis[108] routes.

2. We limit the number of solid reactants to a maximum of $n-1$ where n is the number of chemical species in our target phase (e.g. maximum three solid reactants considered for YBCO synthesis).

3. We consider only two-step pathways; i.e. reaction tree is allowed to be two layers deep. In other words, all intermediates that are used as reactants in reactions leading to target need to be synthesized exclusively from *known* precursors (no further intermediates allowed).

4. We limit the number of first-layer reactions (leading to target) to 10, and that of second-layer reactions (leading to intermediates required for first-layer) to 5 for auxiliary synthesis of each intermediate. These limits ensure obtaining a practical reaction

tree that is not too wide. To rank and select such feasible subsets of reactions based on nucleation potency and phase competition metrics, we use a common approach for multi-objective decision making called TOPSIS.[109]

The reaction tree obtained for solid-state synthesis of YBCO based on these guidelines is shown in Fig. S9 and the respective reactions are listed in Table 5. The top first-layer reactions listed in this tree are obtained from the ranking of ∼530 unique synthesis reactions generated from the given precursor library for YBCO (of these ∼280 involve the shown intermediates). In effect, out of thousands of possible pathways that become available through combining such first-layer reactions with routes for intermediates themselves, this reaction tree covers 42 routes: one of which is a single step pathway and the rest (41 routes) are two-step synthesis pathways. These two-step pathways have first-layer reactions that utilize one of the ternary intermediates $BaCuO_2$, Ba_2CuO_3, $Ba(CuO)_2$ and $Y_2Cu_2O_5$ except. The only one-step route is a peroxide one, BaO_2-Y_2O_3-Cu_2O (20674:R0). Many intermediate phases in fact evolve during synthesis of YBCO,[68,78–80] and hence deliberately utilizing them can help constrain the reactions to follow a desired pathway. Besides, such ternary intermediates require two solid reactants to make in all high-ranking reactions in Table 5, and can therefore be produced at a single reaction interface more controllably. For instance, 20674:R1 uses $BaCuO_2$, which is known to form easily during YBCO synthesis or explicitly from precursors CuO and $BaCO_3$,[68,77] a high-ranking reaction in Table 5 for $BaCuO_2$ (997034:R1). Hence, a two-step progression like 997034:R1⇒20674:R1 is quite plausible for YBCO synthesis. In fact, an almost identical pathway has been reported by Ruckenstein et al.[68] (only difference being use of CuO instead of $0.5\,Cu_2O$, which has negligible effect on reaction energy near temperatures close to CuO→Cu_2O decomposition). The $BaCO_3$ version of first-layer reaction 20674:R7 was also reported to yield YBCO.[68] The intermediate required for the first layer reaction 20674:R3 or 20674:R4; i.e., the Ba_2CuO_3 phase is reported to be amenable to synthesis via the peroxide route BaO_2-CuO (8790:R2) already ranked as favorable by the system,[110,111] implying such two-step pathways are plausible. Overall, experimental findings

in the literature for YBCO provide a certain degree of validation and plausibility for the retrosynthetic pathways obtained via PIRO and the practical strategy we outlined.

Further analyses of the selectivity metric

The case studies, in particular those where the success (e.g. phase purity) of two reactions could be compared, provide empirical confirmation that the competition metric in PIRO functions as intended in ordering reactions by their relative selectivities. To the best of our knowledge, there exists no experimental report that would allow selectivity comparisons across broader reactant sets. But a complementary approach would be *synthesizing* such a dataset, by collating literature reports on formation of *different* products from the reactant sets of a given target, which would serve as a reasonable experimental surrogate to compare the competing phase count metric. Using the dataset of Kononova et al.,[10] we were able to extract several informative targets (Table S1) where the syntheses data available on their *competing phases* were broad enough to obtain a ranking of the selectivity of reactant sets. We find that ordering of reactant sets in PIRO is in good agreement with data extracted from the mined recipes, providing additional validation.

We also investigate the variability in rank ordering of competition metrics of synthesis reactions when different pools of competing phases are available (Fig. S10), using BTO, LCO and YBCO as examples. These pools include all MP entries (i.e. including both ICSD-sourced and hypothetical compounds), a random 50% sub-sampling of MP entries (to emulate the effect of incomplete phase data) and only ICSD-sourced MP entries, and varying *depths* of parasitic reactions (i.e. including phases from sub-spaces). We confirm that all scenarios rank reactions similarly and in line with the ICSD-only pools for YBCO and BTO, where there exists enough data to make a comparison. Supported by this observation, we confirm that using all relevant MP data and/or increasing the competing reaction depth are essential strategies to get informative selectivities for targets where the parent chemical space may not contain enough ICSD entries to get a proper selectivity ranking of reactions

(e.g. even for LCO). Overall, we conclude that including all data available to us in the DFT database, regardless of the origin, in the competition pool is a reliable option to assess selectivity of reactions, as also empirically confirmed in all previous examples showcased.

Current limitations and future work

As in any new predictive method, we should reiterate the basic assumptions/limitations and provide a roadmap for improvement of PIRO. By design the main function of the framework is to distinguish reactions favorable for preferential nucleation of a target phase from those that we cannot make such an assessment within the boundaries of our framework. Reactions away from the Pareto optimality may still be functional in practice, either because they may proceed through a mechanism not modeled well by a CNT-based approach or because of the approximations we make to estimate physical and chemical quantities or render the CNT tractable for high-throughput data. Many of these limitations can be systematically tackled in the near future. For example, Fig. S1 shows that the approximation to capture the deviation of the surface energy ratios from unity via a simple similarity distance metric (which in turn helped simplify Eq.3 to a tractable level in Eq.4) functions exactly as intended in the framework towards isolating reactions with high nucleation potency from those with *unknown* potency (where the ratios spread out with distance). Computed or predicted surface/interface energies, if they become available, can help utilize Eq.3 directly and make more refined predictions, including within the latter class of reactions. One step further, when surfaces/interfaces reconstruct or have kinks, ledges, dislocations or point defects, current predictions deduced from a low-indexed matching of lattice planes may serve as an upper bound in terms of the potential catalytic gains in nucleation.[25,112] Hence, there is an opportunity to extend the current formulation to a broader class of surface imperfections that can catalyze nucleation of a target phase selectively. Temperature effects on Gibbs free energy of solids (neglected here due to computational cost) could be incorporated to obtain more accurate reaction energies, possibly via machine learning predictions.[113] DFT energies for solids

can be improved through use of new functionals in high-throughput datasets.[114] Thermodynamic models can be extended to other synthesis processes such as metathesis reactions. The retrosynthetic planning strategy can be refined using graph-theoretical approaches to chemical reaction networks.[115] Another opportunity is to couple PIRO with external data sources on other properties of *buyable* reagents, and provide guidance on critical experiment design aspects that are not easy to account for within the current level of theory, for example, for how temperature, starting stoichiometries or atmospheric conditions should be adjusted given the excessive volatilities of certain solid reactants (e.g. MoO_3, RuO_2 etc.).

As noted by Jansen and Schön two decades ago,[19,116] an essential step in planning the synthesis of inorganics is *a priori* mapping of the energy landscape, which in our framework would improve the informativeness of the phase competition metric (e.g. especially in under-explored chemistries, and pave the way for more effective planning of synthesis, as we demonstrated also in the previous section. Today, this mapping can be achieved by efficiently populating the space with low-energy hypothetical compounds generated using data-driven methods.[17,117–119] While the current implementation uses MP as the main data source,[29] extensions of the framework to work with other high-throughput DFT databases is also straightforward.[120–123]

Finally, we should stress that while we have demonstrated that PIRO can provide guidance as an effective solid-synthesis planning tool for a wide array of materials, scientists would be the ultimate decision makers in reaction selection, considering many additional factors from safety or toxicity to decomposition temperatures, reactivities, volatilities, instabilities, moisture sensitivities, equipment constraints, and more critically, their field expertise and heuristics.

Conclusion

In summary, we presented a data-driven approach to classical nucleation theory that takes into account structural, chemical and thermodynamic information for identifying catalytically-favorable and phase-selective solid-state synthesis routes for target inorganic crystalline compounds. We tested and validated this computational approach through analyses of the experimental synthesis reports for a diverse array of compounds, and further showed how it enables *retrosynthetic analysis* for designing multi-step synthesis reactions for inorganics.

Acknowledgement

The authors thank Dr. Santosh Suram for fruitful discussions. Sample structure visual in Fig.1 is generated using VESTA.[124]

References

(1) Kanatzidis, M. G. et al. Report from the third workshop on future directions of solid-state chemistry: The status of solid-state chemistry and its impact in the physical sciences. *Progress in Solid State Chemistry* **2008**, *36*, 1–133.

(2) Kohlmann, H. Looking into the Black Box of Solid-State Synthesis. *European Journal of Inorganic Chemistry* **2019**, *2019*, 4174–4180.

(3) Chamorro, J. R.; McQueen, T. M. Progress toward Solid State Synthesis by Design. *Accounts of Chemical Research* **2018**, *51*, 2918–2925.

(4) Martinolich, A. J.; Neilson, J. R. Toward Reaction-by-Design: Achieving Kinetic Control of Solid State Chemistry with Metathesis. *Chemistry of Materials* **2017**, *29*, 479–489.

(5) Brown, W.; Dollimore, D.; Galwey, A. In *Comprehensive Chemical Kinetics*; Bamford, C., Tipper, C., Eds.; 1980; Vol. 22; Chapter 1, pp 1–15.

(6) Turnbull, D.; Fisher, J. C. Rate of nucleation in condensed systems. *The Journal of Chemical Physics* **1949**, *17*, 71–73.

(7) Corey, E. J. The Logic of Chemical Synthesis: Multistep Synthesis of Complex Carbogenic Molecules(Nobel Lecture). *Angewandte Chemie International Edition in English* **1991**, *30*, 455–465.

(8) Kim, E.; Huang, K.; Saunders, A.; McCallum, A.; Ceder, G.; Olivetti, E. Materials Synthesis Insights from Scientific Literature via Text Extraction and Machine Learning. *Chemistry of Materials* **2017**, *29*, 9436–9444.

(9) Kim, E.; Huang, K.; Kononova, O.; Ceder, G.; Olivetti, E. Distilling a Materials Synthesis Ontology. *Matter* **2019**, *1*, 8–12.

(10) Kononova, O.; Huo, H.; He, T.; Rong, Z.; Botari, T.; Sun, W.; Tshitoyan, V.; Ceder, G. Text-mined dataset of inorganic materials synthesis recipes. *Scientific data* **2019**, *6*, 203.

(11) Oliynyk, A. O.; Antono, E.; Sparks, T. D.; Ghadbeigi, L.; Gaultois, M. W.; Meredig, B.; Mar, A. High-Throughput Machine-Learning-Driven Synthesis of Full-Heusler Compounds. *Chemistry of Materials* **2016**, *28*, 7324–7331.

(12) Aykol, M.; Hegde, V. I.; Hung, L.; Suram, S.; Herring, P.; Wolverton, C.; Hummelshøj, J. S. Network analysis of synthesizable materials discovery. *Nature Communications* **2019**, *10*, 2018.

(13) Malik, S. A.; Goodall, R. E. A.; Lee, A. A. Materials Graph Transformer predicts the outcomes of inorganic reactions with reliable uncertainties. **2020**, arXiv:2007.15752.

(14) Jang, J.; Gu, G. H.; Noh, J.; Kim, J.; Jung, Y. Structure-Based Synthesizability Prediction of Crystals Using Partially Supervised Learning. *Journal of the American Chemical Society* **2020**, *142*, 18836–18843.

(15) Sun, W.; Dacek, S. T.; Ong, S. P.; Hautier, G.; Jain, A.; Richards, W. D.; Gamst, A. C.; Persson, K. A.; Ceder, G. The thermodynamic scale of inorganic crystalline metastability. *Science Advances* **2016**, *2*, e1600225–e1600225.

(16) Aykol, M.; Dwaraknath, S. S.; Sun, W.; Persson, K. A. Thermodynamic limit for synthesis of metastable inorganic materials. *Science Advances* **2018**, *4*, eaaq0148.

(17) Sun, W.; Bartel, C. J.; Arca, E.; Bauers, S. R.; Matthews, B.; Orvañanos, B.; Chen, B. R.; Toney, M. F.; Schelhas, L. T.; Tumas, W.; Tate, J.; Zakutayev, A.; Lany, S.; Holder, A. M.; Ceder, G. A map of the inorganic ternary metal nitrides. *Nature Materials* **2019**, *18*, 732–739.

(18) Bianchini, M.; Wang, J.; Clément, R. J.; Ouyang, B.; Xiao, P.; Kitchaev, D.; Shi, T.; Zhang, Y.; Wang, Y.; Kim, H.; Zhang, M.; Bai, J.; Wang, F.; Sun, W.; Ceder, G. The interplay between thermodynamics and kinetics in the solid-state synthesis of layered oxides. *Nature Materials* **2020**,

(19) Jansen, M. A concept for synthesis planning in solid-state chemistry. *Angewandte Chemie - International Edition* **2002**, *41*, 3746–3766.

(20) Brown, W.; Dollimore, D.; Galwey, A. In *Comprehensive Chemical Kinetics*; Bamford, C., Tipper, C., Eds.; 1980; Vol. 22; Chapter 3, pp 41–113.

(21) Brown, W.; Dollimore, D.; Galwey, A. In *Comprehensive Chemical Kinetics*; Bamford, C., Tipper, C., Eds.; 1980; Vol. 22; Chapter 5, pp 247–282.

(22) Khawam, A.; Flanagan, D. R. Solid-state kinetic models: Basics and mathematical fundamentals. *Journal of Physical Chemistry B* **2006**, *110*, 17315–17328.

(23) Turnbull, D.; Vonnegut, B. Nucleation Catalysis. *Industrial Engineering Chemistry* **1952**, *44*, 1292–1298.

(24) Becker, R.; Döring, W. Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. *Annalen der Physik* **1935**, *416*, 719–752.

(25) Russell, K. C. Nucleation in solids: The induction and steady state effects. *Advances in Colloid and Interface Science* **1980**, *13*, 205–318.

(26) Kalikmanov, V. *Nucleation Theory*; 2013.

(27) Boldyrev, V. V. Topochemistry and topochemical reactions. *Reactivity of Solids* **1990**, *8*, 231–246.

(28) Merkle, R.; Maier, J. On the Tammann-rule. *Zeitschrift fur Anorganische und Allgemeine Chemie* **2005**, *631*, 1163–1166.

(29) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials* **2013**, *1*, 011002.

(30) Barin, I. *Thermochemical Data of Pure Substances*, 3rd ed.; VCH, 1995.

(31) Wagman, D.; Evans, W.; Parker, V.; Schumm, R.; Halow, I.; Bailey, S.; Churney, K.; Nuttall, R. The NBS Tables of Chemical of Chemical Thermodynamic Properties. *J. Phys. Chem. Ref. Data* **1982**, *Vol. 11*, Suppl. 2.

(32) Wang, L.; Maxisch, T.; Ceder, G. Oxidation energies of transition metal oxides within the GGA+U framework. *Phys. Rev. B* **2006**, *73*, 195107.

(33) Grindy, S.; Meredig, B.; Kirklin, S.; Saal, J.; Wolverton, C. Approaching chemical accuracy with density functional calculations : Diatomic energy corrections. *Physical Review B* **2013**, *87*, 075150.

(34) Zur, A.; McGill, T. C. Lattice match: An application to heteroepitaxy. *Journal of Applied Physics* **1984**, *55*, 378–386.

(35) Ding, H.; Dwaraknath, S. S.; Garten, L.; Ndione, P.; Ginley, D.; Persson, K. A. Computational Approach for Epitaxial Polymorph Stabilization through Substrate Selection. *ACS Applied Materials and Interfaces* **2016**, *8*, 13086–13093.

(36) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science* **2013**, *68*, 314–319.

(37) Ward, L.; Liu, R.; Krishna, A.; Hegde, V. I.; Agrawal, A.; Choudhary, A.; Wolverton, C. Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations. *Physical Review B* **2017**, *96*, 024104.

(38) Ward, L. et al. Matminer: An open source toolkit for materials data mining. *Computational Materials Science* **2018**, *152*, 60–69.

(39) Belsky, A.; Hellenbrandt, M.; Karen, V. L.; Luksch, P. New developments in the Inorganic Crystal Structure Database (ICSD): accessibility in support of materials research and design. *Acta Crystallographica Section B Structural Science* **2002**, *58*, 364–369.

(40) Zhang, Y.; Furness, J. W.; Xiao, B.; Sun, J. Subtlety of TiO2 phase stability: Reliability of the density functional theory predictions and persistence of the self-interaction error. *Journal of Chemical Physics* **2019**, *150*.

(41) Vijatović, M. M.; Bobić, J. D.; Stojanović, B. D. History and challenges of barium titanate: Part I. *Science of Sintering* **2008**, *40*, 155–165.

(42) Stojanovic, B. D.; Simoes, A. Z.; Paiva-Santos, C. O.; Jovalekic, C.; Mitic, V. V.; Varela, J. A. Mechanochemical synthesis of barium titanate. *Journal of the European Ceramic Society* **2005**, *25*, 1985–1989.

(43) Cournil, M.; Soustelle, M.; Thomas, G. Solid-solid reactions, II: Mechanism of barium metatitanate synthesis. *Oxidation of Metals* **1979**, *13*, 89–104.

(44) Buscaglia, M. T.; Bassoli, M.; Buscaglia, V.; Alessio, R. Solid-state synthesis of ultrafine BaTiO 3 powders from nanocrystalline BaCO 3 and TiO 2. *Journal of the American Ceramic Society* **2005**, *88*, 2374–2379.

(45) Beauger, A.; Mutin, J. C.; Niepce, J. C. Synthesis reaction of metatitanate BaTiO3 - Part 2 Study of solid-solid reaction interfaces. *Journal of Materials Science* **1983**, *18*, 3543–3550.

(46) Lee, T. T.; Huang, C. Y.; Chang, C. Y.; Lin, S. P.; Su, C. Y.; Lee, C. T.; Fujimoto, M. Phase evolution and nucleus growth observation of solid-state BaTiO 3 powder prepared by high-energy bead milling for raw material mixing. *Japanese Journal of Applied Physics* **2011**, *50*.

(47) Komarov, A. V.; Parkin, I. P. New routes in the self-propagating high-temperature synthesis of barium titanium oxide. *Polyhedron* **1996**, *15*, 1349–1353.

(48) Licheri, R.; Fadda, S.; Orrù, R.; Cao, G.; Buscaglia, V. Self-propagating high-temperature synthesis of barium titanate and subsequent densification by spark plasma sintering (SPS). *Journal of the European Ceramic Society* **2007**, *27*, 2245–2253.

(49) Larson, E. M.; Wong, J.; Holt, J. B.; Waide, P. A.; Rupp, B. The combustion synthesis of the ferroelectric material, BaTiO3, studied by time-resolved X-ray diffraction. *Powder Diffraction* **1999**, *14*, 111–113.

(50) Lotnyk, A.; Senz, S.; Hesse, D. Formation of BaTiO3 thin films from (110) TiO2 rutile single crystals and BaCO3 by solid state reactions. *Solid State Ionics* **2006**, *177*, 429–436.

(51) Rössel, M.; Gablenz, S.; Müller, T.; Röder, A.; Abicht, H. P. A core-shell structured BaTiO3 precursor preparation, characterization and potential. *Analytical and Bioanalytical Chemistry* **2003**, *375*, 310–314.

(52) Rössel, M.; Höche, H. R.; Leipner, H. S.; Völtzke, D.; Abicht, H. P.; Hollricher, O.; Müller, J.; Gablenz, S. Raman microscopic investigations of BaTiO3 precursors with core-shell structure. *Analytical and Bioanalytical Chemistry* **2004**, *380*, 157–162.

(53) Beauger, A.; Mutin, J. C.; Niepce, J. C. Synthesis reaction of metatitanate BaTiO3 - Part 1 Effect of the gaseous atmosphere upon the thermal evolution of the system BaCO3-TiO2. *Journal of Materials Science* **1983**, *18*, 3041–3046.

(54) Pfaff, G. Synthesis and characterization of Ba2TiO4. *Journal of Materials Science Letters* **1991**, *10*, 129–131.

(55) Zhu, N.; West, A. R. Formation and stability of ferroelectric BaTi2O5. *Journal of the American Ceramic Society* **2010**, *93*, 295–300.

(56) Mizushima, K.; Jones, P. C.; Wiseman, P. J.; Goodenough, J. B. LixCoO2 (0<x<1): A new cathode material for batteries of high energy density. *Solid State Ionics* **1981**, *3-4*, 171–174.

(57) Antolini, E. Preparation and properties of Li-Co-O compounds. *Journal of the European Ceramic Society* **1998**, *18*, 1405–1411.

(58) Antolini, E. LiCoO2: Formation, structure, lithium and oxygen nonstoichiometry, electrochemical behaviour and transport properties. *Solid State Ionics* **2004**, *170*, 159–171.

(59) Lundblad, A.; Bergman, B. Synthesis of LiCoO2 starting from carbonate precursors I. The reaction mechanisms. *Solid State Ionics* **1997**, *96*, 173–181.

(60) Carewska, M.; Di Bartolomeo, A.; Scaccia, S. Thermoanalytical study of synthesis of mixed lithium cobalt oxides. *Thermochimica Acta* **1995**, *269-270*, 491–506.

(61) Berbenni, V.; Milanese, C.; Bruni, G.; Marini, A. Solid state synthesis of stoichiometric LiCoO2 from mechanically activated Co-Li2CO3 mixtures. *Materials Chemistry and Physics* **2006**, *100*, 251–256.

(62) Johnston, W. D.; Heikes, R. R.; Sestrich, D. The preparation, crystallography, and magnetic properties of the LixCo(1-x)O system. *Journal of Physics and Chemistry of Solids* **1958**, *7*, 1–13.

(63) Kobayashi, H.; Tsukasaki, T.; Ogasawara, Y.; Hibino, M.; Kudo, T.; Mizuno, N.; Honma, I.; Yamaguchi, K. Cation-Disorder-Assisted Reversible Topotactic Phase Transition between Antifluorite and Rocksalt Toward High-Capacity Lithium-Ion Batteries. *ACS Applied Materials and Interfaces* **2020**, *12*, 43605–43613.

(64) Möller, A. Synthesis, Structures, Magnetic Properties, and Absorption Spectra of the Alkaline Oxocobaltates(II): Li6CoO4, Na4CoO3, and Na10Co4O9. *Chemistry of Materials* **1998**, *10*, 3196–3201.

(65) Amatucci, G.; Tarascon, J.; Klein, L. CoO2, The End Member of the Li x CoO2 Solid Solution. *J. Electrochem. Soc.* **1996**, *143*, 1114–1123.

(66) Wu, M. K.; Ashburn, J. R.; Torng, C. J.; Hor, P. H.; Meng, R. L.; Gao, L.; Huang, Z. J.; Wang, Y. Q.; Chu, C. W. Superconductivity at 93 K in a new mixed-phase Yb-Ba-Cu-O compound system at ambient pressure. *Physical Review Letters* **1987**, *58*, 908–910.

(67) Sugai, T.; Oya, G.-i.; Imai, S. Thermoanalytical Studies on Preparation Conditions of

Superconducting YBa 2 Cu 3 O 7- y. *Japanese Journal of Applied Physics* **1989**, *28*, 341–345.

(68) Ruckenstein, E.; Narain, S.; Wu, N.-L. Reaction pathways for the formation of the YBa2Cu3O7 compound. *Journal of Materials Research* **1989**, *4*, 267–272.

(69) Pathak, L. C.; Mishra, S. K. A review on the synthesis of Y-Ba-Cu-oxide powder. *Superconductor Science and Technology* **2005**, *18*.

(70) Fahlman, B. D. Superconductor Synthesis - An Improvement. *Journal of Chemical Education* **2001**, *78*, 1182.

(71) Lebrat, J. P.; Varma, A. Some Further Studies in Combustion Synthesis of the YBa2 Cu3 O7-x Superconductor. *Combustion Science and Technology* **1993**, *88*, 177–185.

(72) Kao, M. Synthesis of Y-Ba-Cu-O superconductors from Y2O3 BaO2, and Cu2O: The optional oxygen treatment. *Materials Letters* **1987**, *6*, 53–57.

(73) Kao, M.; McKinney, B. The properties of YBa2Cu3O7 superconductors sintered from the powder mixture of Y2O3, BaO2, and Cu2O. *Materials Letters* **1991**, *11*, 91–95.

(74) Hepp, A.; Gaier, J. Inert atmosphere preparation of Ba2YCu3O7 using BaO2. *Materials Research Bulletin* **1988**, *23*, 693–700.

(75) Costa, C. A.; Ferretti, M.; Olcese, C. L.; Cimberle, M. R.; Ferdeghini, C.; Nicchiotti, G. L.; Siri, A. S.; Rizzuto, C. Synthesis of YBa2Cu3O7-x polycrystalline superconductors from Ba peroxide: First physico-chemical characterization. *Journal of Crystal Growth* **1987**, *85*, 623–627.

(76) Miura, A.; Bartel, C. J.; Goto, Y.; Mizuguchi, Y.; Moriyoshi, C.; Kuroiwa, Y.; Wang, Y.; Yaguchi, T.; Shirai, M.; Nagao, M.; Rosero-Navarro, N. C.; Tadanaga, K.; Ceder, G.; Sun, W. Sequential pairwise reactions dictate phase evolution in the solid-state synthesis of multicomponent ceramics. **2020**, arXiv:2009.10896.

(77) Flor, G.; Scavini, M.; Anselmi-Tamburini, U.; Spinolo, G. Kinetics and mechanisms of the formation of YBa2Cu3O7-x. *Solid State Ionics* **1990**, *43*, 77–83.

(78) Peterson, D. E.; Kubat-Martin, K. A.; George, T. G.; Zocco, T. G.; Thompson, J. D. Synthesis of YBa 2 Cu 3 O 7âĹŠ x by chemical precursors. *Journal of Materials Research* **1991**, *6*, 11–17.

(79) Cogdell, C.; Wayment, D.; Kubat-Martin, K. A Convenient, One-Step Synthesis of YBa2CU3O7-x Superconductors: An Undergraduate Inorganic/Materials Laboratory Experiment. *J. Chem. Educ.* **1995**, *72*, 840–841.

(80) Anselmi-Tamburini, U.; Ghigna, P.; Spinolo, G.; Flor, G. Solid state synthesis of YBa2Cu3O7âĹŠx: Mechanisms of BaCuO2 formation. *Journal of Physics and Chemistry of Solids* **1991**, *52*, 715–721.

(81) Zhou, Q.; Kennedy, B. J.; Kimpton, J. A. The effect of disorder in Ba2YTaO6 on the tetragonal to cubic phase transition. *Journal of Solid State Chemistry* **2011**, *184*, 729–734.

(82) Mouyane, M.; Womes, M.; Jumas, J. C.; Olivier-Fourcade, J.; Lippens, P. E. Original electrochemical mechanisms of CaSnO3 and CaSnSiO 5 as anode materials for Li-ion batteries. *Journal of Solid State Chemistry* **2011**, *184*, 2877–2886.

(83) Zhang, W.; Yu, H.; Cantwell, J.; Wu, H.; Poeppelmeier, K. R.; Halasyamani, P. S. LiNa5Mo9O30: Crystal Growth, Linear, and Nonlinear Optical Properties. *Chemistry of Materials* **2016**, *28*, 4483–4491.

(84) Blum, C. G.; Holcombe, A.; Gellesch, M.; Sturza, M. I.; Rodan, S.; Morrow, R.; Maljuk, A.; Woodward, P.; Morris, P.; Wolter, A. U.; Büchner, B.; Wurmehl, S. Flux growth and characterization of Sr2NiWO6 single crystals. *Journal of Crystal Growth* **2015**, *421*, 39–44.

(85) Artini, C.; Carnasciali, M. M.; Costa, G. A.; Ferretti, M.; Cimberle, M. R.; Putti, M.; Masini, R. Synthesis and characterisation of superconducting RuSr2GdCu2O8. *Physica C* **2002**, *377*, 431–436.

(86) Bridges, C. A.; Greedan, J. E. Phase transitions and electrical transport in the mixed-valence V 2+/V3+ oxide BaV10O15. *Journal of Solid State Chemistry* **2004**, *177*, 1098–1110.

(87) Lufaso, M. W.; Gemmill, W. R.; Mugavero, S. J.; Lee, Y.; Vogt, T.; zur Loye, H. C. Compression mechanisms of symmetric and Jahn-Teller distorted octahedra in double perovskites: A2CuWO6 (A=Sr, Ba), Sr2CoMoO6, and La2LiRuO6. *Journal of Solid State Chemistry* **2006**, *179*, 3556–3561.

(88) Kimani, M. M.; McMillen, C. D.; Kolis, J. W. Synthetic and spectroscopic studies of vanadate glaserites II: Photoluminescence studies of Ln:K3Y(VO4)2 (Ln=Eu, Er, Sm, Ho, or Tm). *Journal of Solid State Chemistry* **2015**, *226*, 320–325.

(89) Chi, E. O.; Ok, K. M.; Porter, Y.; Halasyamani, P. S. Na 2Te 3Mo 3O 16: A new molybdenum tellurite with second-harmonic generating and pyroelectric properties. *Chemistry of Materials* **2006**, *18*, 2070–2074.

(90) Picard, S.; Gougeon, P.; Potel, M. Synthesis, Structural Evolution, and Theoretical and Physical Studies of the Novel Compounds M 2 Mo 9 S 11 (M = K, Rb) and Related Metastable Materials Cu x K 1.8 Mo 9 S 11 (x = 0 or 2) Containing Bioctahedral Mo 9 Clusters. *Inorganic Chemistry* **1999**, *38*, 4422–4429.

(91) Hannerz, H.; Svensson, G.; Istomin, S. Y.; D'Yachenko, O. G. Transmission Electron Microscopy and Neutron Powder Diffraction Studies of GdFeO3 Type SrNbO3. *Journal of Solid State Chemistry* **1999**, *147*, 421–428.

(92) MacQuart, R. B.; Kennedy, B. J.; Avdeev, M. A primitive tetragonal intermedi-

ate in the orthorhombiccubic phase transition of perovskite-type strontium niobate Sr0.92NbO3. *Journal of Solid State Chemistry* **2010**, *183*, 2400–2405.

(93) Inami, T. Neutron powder diffraction experiments on the layered triangular-lattice antiferromagnets RbFe(MoO4)2 and CsFe(SO4)2. *Journal of Solid State Chemistry* **2007**, *180*, 2075–2079.

(94) Lii, K. H.; Tsai, H. J. K2V3P4O17: A Vanadium (IV) Pyrophosphate with a Layer Structure. *Journal of Solid State Chemistry* **1990**, *401*, 396–401.

(95) Bräuchle, S.; Huppertz, H. Synthesis and structural characterization of the new rare-earth borosilicates Pr3BSi2O10 and Tb3BSi2O10. *Zeitschrift fur Naturforschung - Section B Journal of Chemical Sciences* **2015**, *70*, 929–934.

(96) Wu, S. P.; Chen, D. F.; Jiang, C.; Mei, Y. X.; Ma, Q. Synthesis of monoclinic CaSnSiO5 ceramics and their microwave dielectric properties. *Materials Letters* **2013**, *91*, 239–241.

(97) Du, X.; Gao, Z.; Liu, F.; Guo, X.; Wang, X.; Sun, Y.; Tao, X. Anisotropic properties and Raman spectra of a LiNa 5 Mo 9 O 30 single crystal grown by the TSSG method. *CrystEngComm* **2020**, *22*, 7716–7722.

(98) Tanwar, K.; Saxena, M.; Maiti, T. Enhancement of thermoelectric power factor of Sr2CoMoO6 double perovskite by annealing in reducing atmosphere. *Journal of Applied Physics* **2017**, *122*.

(99) Bauernfeind, L.; Widder, W.; Braun, H. F. Ruthenium-based layered cuprates RuSr2LnCu2O8 and RuSr2(Ln1+xCe1-x)Cu2O10 (LnSm, Eu and Gd). *Physica C: Superconductivity and its applications* **1995**, *254*, 151–158.

(100) Ridgley, D.; Ward, R. The Preparation of a Strontium-Niobium Bronze with the Perovskite Structure 1. *Journal of the American Chemical Society* **1955**, *77*, 6132–6136.

(101) Liu, G.; Greedan, J. E. Structure and magnetism in AB10O15+X (A = Sr, Ba; B = V, Cr; x = 0, âĹij1): Evidence for geometric frustration. *Journal of Solid State Chemistry* **1996**, *122*, 416–427.

(102) Retuerto, M.; Alonso, J. A.; García-Hernández, M.; Martínez-Lope, M. J. Synthesis, structure and magnetic properties of the new double perovskite Ca2CrSbO6. *Solid State Communications* **2006**, *139*, 19–22.

(103) Galasso, F.; Darby, W. Preparation and properties of Sr2FeO3F. *Journal of Physical Chemistry* **1963**, *67*, 1451–1453.

(104) Akimoto, J.; Takei, H. Synthesis and crystal structure of NaTi8O13. *Journal of Solid State Chemistry* **1991**, *90*, 147–154.

(105) Vennos, D. A.; DiSalvo, F. J. Synthesis and characterization of a new ternary nitride, Ca3VN3. *Journal of Solid State Chemistry* **1992**, *98*, 318–322.

(106) Coley, C. W.; Rogers, L.; Green, W. H.; Jensen, K. F. Computer-Assisted Retrosynthesis Based on Molecular Similarity. *ACS Central Science* **2017**, *3*, 1237–1245.

(107) Martirosyan, K. S.; Luss, D. Carbon combustion synthesis of complex oxides: Process demonstration and features. *AIChE Journal* **2005**, *51*, 2801–2810.

(108) Munir, Z. A.; Anselmi-Tamburini, U. Self-propagating exothermic reactions: The synthesis of high-temperature materials by combustion. *Materials Science Reports* **1989**, *3*, 279–365.

(109) Hwang, C.-L.; Lai, Y.-J.; Liu, T.-Y. A new approach for multiple objective decision making. *Computers Operations Research* **1993**, *20*, 889–899.

(110) Armstrong, A. R.; Janes, R.; Singh, K. K.; Edwards, P. P. Electron spin resonance studies of some cuprate (II) systems. *Bull. Mater. Sci.* **1991**, *14*, 641–649.

(111) Umezawa, A.; Zhang, W.; Gurevich, A.; Feng, Y.; Hellstrom, E. E.; Larbalestier, D. C. Flux pinning, granularity and the irreversibility line of the high-Tc superconductor HgBa2Cu04+x. *Nature* **1993**, *364*, 129–131.

(112) Venables, J. A.; Spiller, G. D.; Hanbucken, M. Nucleation and growth of thin films. *Reports on Progress in Physics* **1984**, *47*, 399–459.

(113) Bartel, C. J.; Millican, S. L.; Deml, A. M.; Rumptz, J. R.; Tumas, W.; Weimer, A. W.; Lany, S.; Stevanović, V.; Musgrave, C. B.; Holder, A. M. Physical descriptor for the Gibbs energy of inorganic crystalline solids and temperature-dependent materials chemistry. *Nature Communications* **2018**, *9*, 4168.

(114) Zhang, Y.; Kitchaev, D. A.; Yang, J.; Chen, T.; Dacek, S. T.; Sarmiento-Pérez, R. A.; Marques, M. A. L.; Peng, H.; Ceder, G.; Perdew, J. P.; Sun, J. Efficient first-principles prediction of solid stability: Towards chemical accuracy. *npj Computational Materials* **2018**, *4*, 9.

(115) McDermott, M.; Dwaraknath, S.; Persson, K. A graph-based network for predicting chemical reaction pathways in solid-state materials synthesis. **2020**, doi:10.21203/rs.3.rs–38000/v1.

(116) Christian Schön, J.; Jansen, M. First step towards planning of syntheses in solid-state chemistry: Determination of promising structure candidates by global optimization. *Angewandte Chemie (International Edition in English)* **1996**, *35*, 1286–1304.

(117) Woodley, S. M.; Catlow, R. Crystal structure prediction from first principles. *Nature materials* **2008**, *7*, 937–46.

(118) Hautier, G.; Fischer, C. C.; Jain, A.; Mueller, T.; Ceder, G. Finding natures missing ternary oxide compounds using machine learning and density functional theory. *Chemistry of Materials* **2010**, *22*, 3762–3767.

(119) Montoya, J. H.; Winther, K. T.; Flores, R. A.; Bligaard, T.; Hummelshøj, J. S.; Aykol, M. Autonomous intelligent agents for accelerated materials discovery. *Chemical Science* **2020**, 8517–8532.

(120) Curtarolo, S.; Setyawan, W.; Wang, S.; Xue, J.; Yang, K.; Taylor, R. H.; Nelson, L. J.; Hart, G. L. W.; Sanvito, S.; Buongiorno-Nardelli, M.; Mingo, N.; Levy, O. AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations. *Computational Materials Science* **2012**, *58*, 227–235.

(121) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. *npj Computational Materials* **2015**, *1*, 15010.

(122) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials design and discovery with high-throughput density functional theory: The Open Quantum Materials Database (OQMD). *JOM* **2013**, *65*, 1501–1509.

(123) Choudhary, K. et al. The joint automated repository for various integrated simulations (JARVIS) for data-driven materials design. *npj Computational Materials* **2020**, *6*, 173.

(124) Momma, K.; Izumi, F. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *Journal of Applied Crystallography* **2011**, *44*, 1272–1276.