

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 $\text{YBa}_2\text{Cu}_3\text{O}_7$, 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 identify the most favorable 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 $\alpha_i \in \{\alpha_1, \alpha_2, \dots, \alpha_n\}$, we write the rate of heterogeneous nucleation of a target phase β

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

$$J_{\alpha_i \rightarrow \beta} = J_0 \cdot \exp[-\Delta G_{\alpha_i \rightarrow \beta}^*/k_B T] \quad (1)$$

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

$$\Delta G_{\alpha_i \rightarrow \beta}^* = (16\pi/3)\gamma_{\beta v}^3[\Delta G_k]^{-2}f(S_{\alpha_i \rightarrow \beta}) \quad (2)$$

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 $f(S_{\alpha_i \rightarrow \beta})$ 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 $S_{\alpha_i \rightarrow \beta}$ (the cosine of the contact angle $\theta_{\beta\alpha_i}$ between substrate α_i and β) described by:

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

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: $(2 - 3S_{\alpha_i \rightarrow \beta} + S_{\alpha_i \rightarrow \beta}^3)/4$.

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 $S_{\alpha_i \rightarrow \beta}$ 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 $f(S_{\alpha_i \rightarrow \beta})$. This condition can simply be achieved in CNT by maximizing $S_{\alpha_i \rightarrow \beta}$ 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. $S_{\alpha_i \rightarrow \beta} \rightarrow 1$. 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 $\gamma_{\beta \alpha_i} / \gamma_{\beta v} \rightarrow 0$ 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 $S_{\alpha_i \rightarrow \beta} = 1$ to capture the deviation from this ideal upper bound as:

$$S_{\alpha_i \rightarrow \beta} \approx 1 - q_{\alpha_i, \beta}^{sim} - q_{\alpha_i, \beta}^{epi} \quad (4)$$

where $q_{\alpha_i, \beta}^{sim}$ and $q_{\alpha_i, \beta}^{epi}$ 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 $q \rightarrow 0$ 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 $\Delta G_{\alpha_i \rightarrow \beta}^*$ 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, $\Delta G_{\alpha_i \rightarrow \beta}^*$ 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 $\Delta E_{\alpha_i \rightarrow \beta}$. Considering that solid-state synthesis almost exclusively takes place at moderately high-temperatures and $\Delta E_{\alpha_i \rightarrow \beta}$ 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 rule²⁸) and use of starting materials that decompose (e.g. carbonates) upon heating. As another option, $\Delta E_{\alpha_i \rightarrow \beta}$ can be used as a parameter deduced from potential diffusion bottlenecks, and lumped together with $\Delta G_{\alpha_i \rightarrow \beta}^*$ to obtain an effective barrier to rank reactions, $\Delta G_{\alpha_i \rightarrow \beta}^{**} \approx \Delta G_{\alpha_i \rightarrow \beta}^* + \Delta E_{\alpha_i \rightarrow \beta}$. 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 $\Delta E_{\alpha_i \rightarrow \beta} \propto q_{\alpha_i, \beta}^{sim}$. Ultimately, for a given candidate reaction k , we assign the minimum catalytic barrier to nucleation we find among reactants, $\Delta G_k^{**} = \min\{\Delta G_{\alpha_1 \rightarrow \beta}^{**}, \Delta G_{\alpha_2 \rightarrow \beta}^{**}, \dots, \Delta G_{\alpha_n \rightarrow \beta}^{**}\}$.

The relative potency of reaction k for catalytic nucleation of a given target β can now be compared through ΔG_k^{**} values, but an assessment of whether the same reaction would yield any alternative (or parasitic) phases other than β by comparison of nucleation rates of different products would require a *quantitative* description of surface/interface energy terms entering Eqs. 2 and 3. Instead, we turn to a simple heuristic that captures the phase competition based on thermodynamics only: the larger the number of other (parasitic) products the reactants of a reaction k can be rebalanced to yield (as the exact ratio of reactants do not have much bearing on the microscopic nucleation process²¹) with a favorable (negative) 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 . This concept is illustrated in Fig. 1, where we see that minimization of this metric directly

competes with ΔG_k , and in effect, the nucleation barrier.

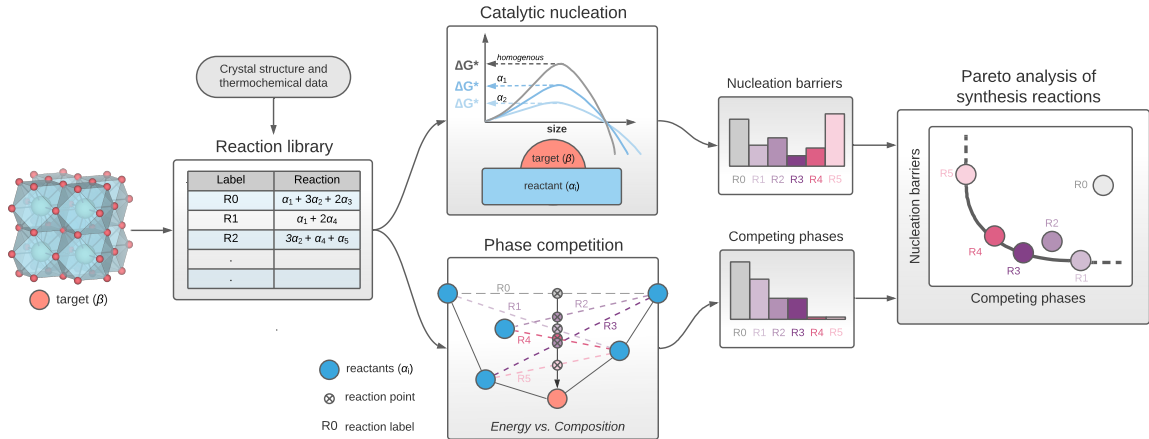


Figure 1: Schematic illustration of the main computational steps in rational solid-state synthesis planning approach, PIRO. Given a target phase β , one begins by building a reaction library enumerating all possible reactions leading to the target phase, then analyzing each of these reactions on the 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.

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 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. O₂, CO, CO₂, H₂O, N₂ etc.) using the experimental standard values obtained from Barin³⁰ 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 CO₃²⁻ (or equivalently per CO₂, 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 10³ Å². 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 MatMiner³⁸ 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 $\Delta G_{\alpha_i \rightarrow \beta}^{**}$, we adopt $\Delta E_{\alpha_i \rightarrow \beta} \approx Cq_{\alpha_i, \beta}^{sim}$, treating C as a hyper-parameter set at 10 eV and assuming $\gamma_{\beta v} \approx 2$ J/m², 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 $\Delta G_{\alpha_i \rightarrow \beta}^{**}$ 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 interactive versions of the recommendation plots reported in this work can be found in Supporting Information. 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 $\text{YBa}_2\text{Cu}_3\text{O}_7$. The vast literature

on synthesis of these technological materials allows us to 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 BaTiO₃. 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, ICSD³⁹ as a proxy). Additionally, we may filter reactant candidates by (iii) whether they are *intermediates* (e.g. ternary compounds other than the target tetragonal BaTiO₃ 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 TiO₂) or peroxides (e.g. BaO₂). 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 BaTiO₃

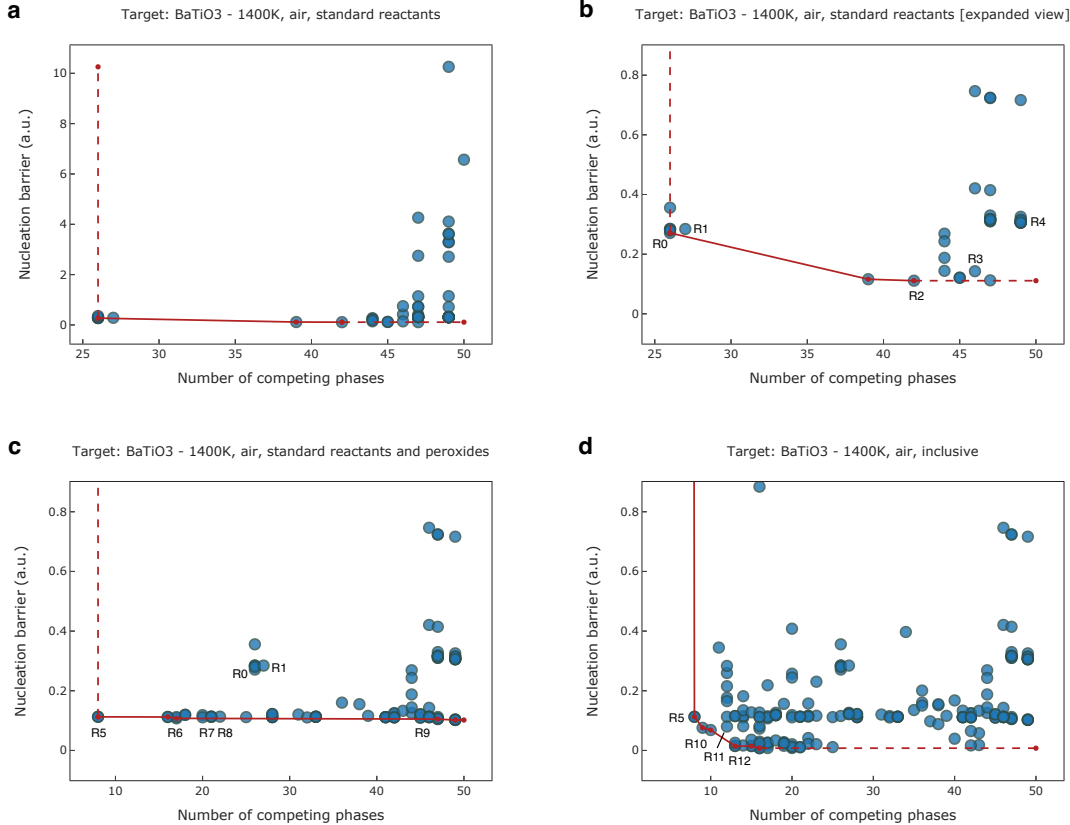


Figure 2: Route planning plots for synthesis of BaTiO₃ under typical thermodynamic conditions using various sets of starting materials. Standard reactants include elements, stable binary compounds and carbonates. The set labeled as "inclusive" contains all stable compounds in the Ba-Ti-O-C system. All 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.

The tetragonal perovskite BaTiO₃ 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 BaCO₃ or BaO, and TiO₂ are used to conventionally make BTO.^{40–45} 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 (BaO₂) have also been reported,^{46–48}

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.^{43–45,49–53} 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.^{44,50–52}), 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 synthesize than BTO⁵⁴ 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.

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.^{55–60} 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 the temperature) which then reacts with the available Li precursor.^{56,58,59} Next, we test

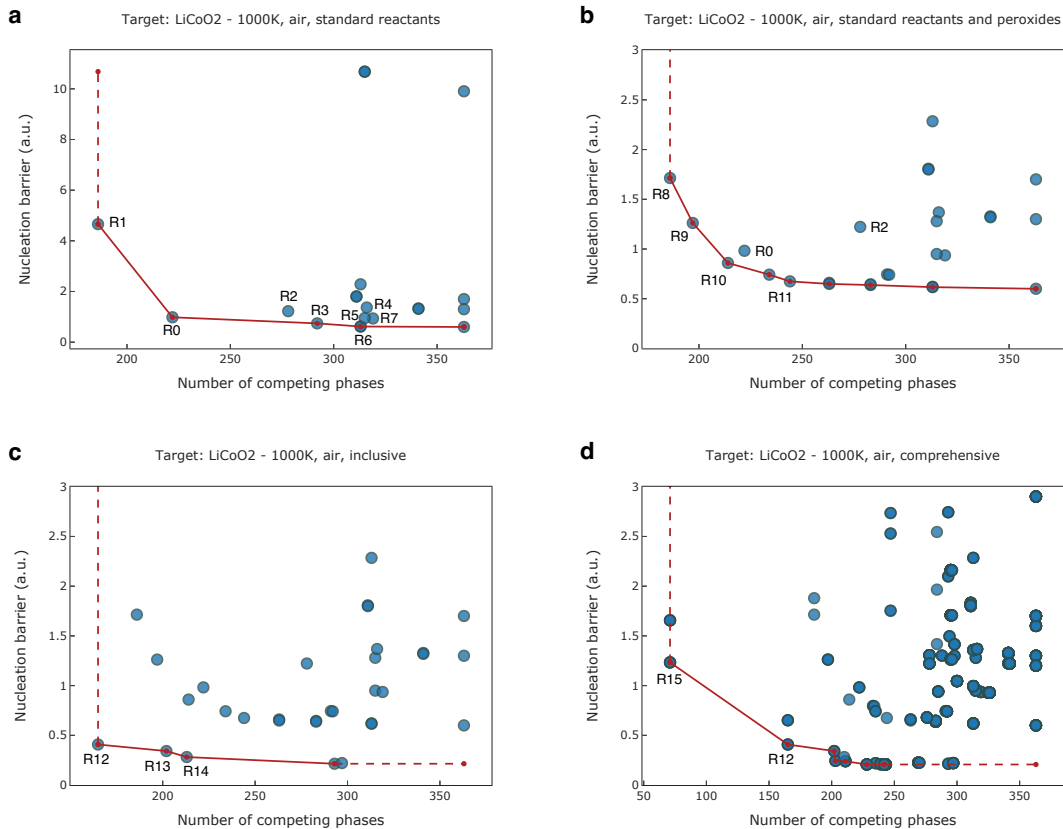


Figure 3: Route planning plots for synthesis of high-temperature (HT) form of layered LiCoO₂ under typical thermodynamic conditions using various sets of starting materials. Standard reactants include elements, stable binary compounds and carbonates. The set labeled as "inclusive" contains all stable 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.

Table 1: Reactions highlighted in Fig.2 for solid-state BaTiO₃ 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 ₂ O ₃ (458) + 0.25 O ₂ (1091399) + 1.0 BaO(1342)
R3	0.5 Ti ₂ O ₃ (458) + 0.25 O ₂ (1091399) + -1.0 CO ₂ + 1.0 BaCO ₃ (5504)
R4	1.0 TiC(631) + 2.0 O ₂ (1091399) + -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 ₂ O ₃ (458) + 0.5 BaO(1342)
R7	1.0 BaO ₂ (1105) + 0.3333 Ti ₂ O ₃ (458) + 0.3333 Ti(72)
R8	1.0 BaO ₂ (1105) + 0.5 Ti ₂ O 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 ₂ O ₃ (458) + 0.3333 BaO ₂ (1105) + 0.3333 Ba ₂ TiO ₄ (3397)
R11	0.5 BaO(1342) + 0.5 BaTi ₂ O ₅ (3943)
R12	0.5 BaTi ₂ O 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.

whether *less conventional* but potentially more favorable routes exist by allowing peroxide reactants (Fig.3(b)) and find that several Li₂O₂ 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 Li₂O₂ and CoO (R8) was in fact reported by Johnston et al.⁶¹ 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 antiferro Li₆CoO₄ bearing reactions (Table 2). The favorability of such reactions (e.g. R12, R13 and R14) in Fig.3(c) is straightforward to reason: Li₆CoO₄ 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.⁶² We find that synthesis of Li₆CoO₄ 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 Li₂O, and is consistent

with experiments.⁶³ 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. Along with the BaTiO₃ example, the confirmed viability of multi-step synthesis routes motivated us to formulate a retrosynthetic reaction planning strategy that uses PIRO, 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: CoO₂ + Li) become part of the Pareto front. While this is not a practical synthesis reaction *per se* (as the practical route to CoO₂ itself is delithiation of LCO⁶⁴), its emergence reinforces the plausibility of the criteria used in the PIRO approach, given that identical metrics; i.e. structural and chemical relationship between the two layered phases and the large chemical driving force are also why LCO has become the workhorse of the battery industry.

Case study: Superconducting YBa₂Cu₃O₇

The chemistry that covers YBa₂Cu₃O₇, 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 Y₂O₃, CuO (or Cu₂O) and BaCO₃ (or BaO),^{65–67} and are found to be Pareto optimal (R0, R1 and R3) in 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). However, 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.⁶⁸ Among these, peroxide (BaO₂) driven combustion routes with different Cu sources, such

Table 2: Reactions highlighted in Fig.3 for solid-state HT-LiCoO₂ 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 ₂ O(1960) + 0.0833 O ₂ + 0.3333 Co ₃ O ₄ (18748)
R1*	1.0 CoO(22408) + 0.5 Li ₂ O(1960) + 0.25 O ₂
R2*	1.0 CoO(22408) + -0.5 CO ₂ + 0.25 O ₂ + 0.5 Li ₂ CO ₃ (3054)
R3*	0.0833 O ₂ + 0.3333 Co ₃ O ₄ (18748) + -0.5 CO ₂ + 0.5 Li ₂ CO ₃ (3054)
R4*	0.5 Li ₂ O(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 ₂ CO ₃ (3054)
R6*	1.0 Li(135) + 0.3333 O ₂ + 0.3333 Co ₃ O ₄ (18748)
R7*	-1.5 CO ₂ + 0.25 O ₂ + 0.5 Li ₂ CO ₃ (3054) + 1.0 CoCO ₃ (21434)
R8*	1.0 CoO(22408) + 0.5 Li ₂ O ₂ (841)
R9**	0.5 Li ₂ O ₂ (841) + -1.0 CO ₂ + 1.0 CoCO ₃ (21434)
R10**	0.3333 Li ₂ O(1960) + 0.1667 Li ₂ O ₂ (841) + 0.3333 Co ₃ O ₄ (18748)
R11**	0.5 Li ₂ O ₂ (841) + 0.25 Co ₃ O ₄ (18748) + 0.25 Co(54)
R12	0.25 O ₂ + 0.8333 CoO(22408) + 0.1667 Li ₆ CoO ₄ (18925)
R13	0.1111 O ₂ + 0.1667 Li ₆ CoO ₄ (18925) + 0.2778 Co ₃ O ₄ (18748)
R14	0.2 Li ₂ O ₂ (841) + 0.1 Li ₆ CoO ₄ (18925) + 0.3 Co ₃ O 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.

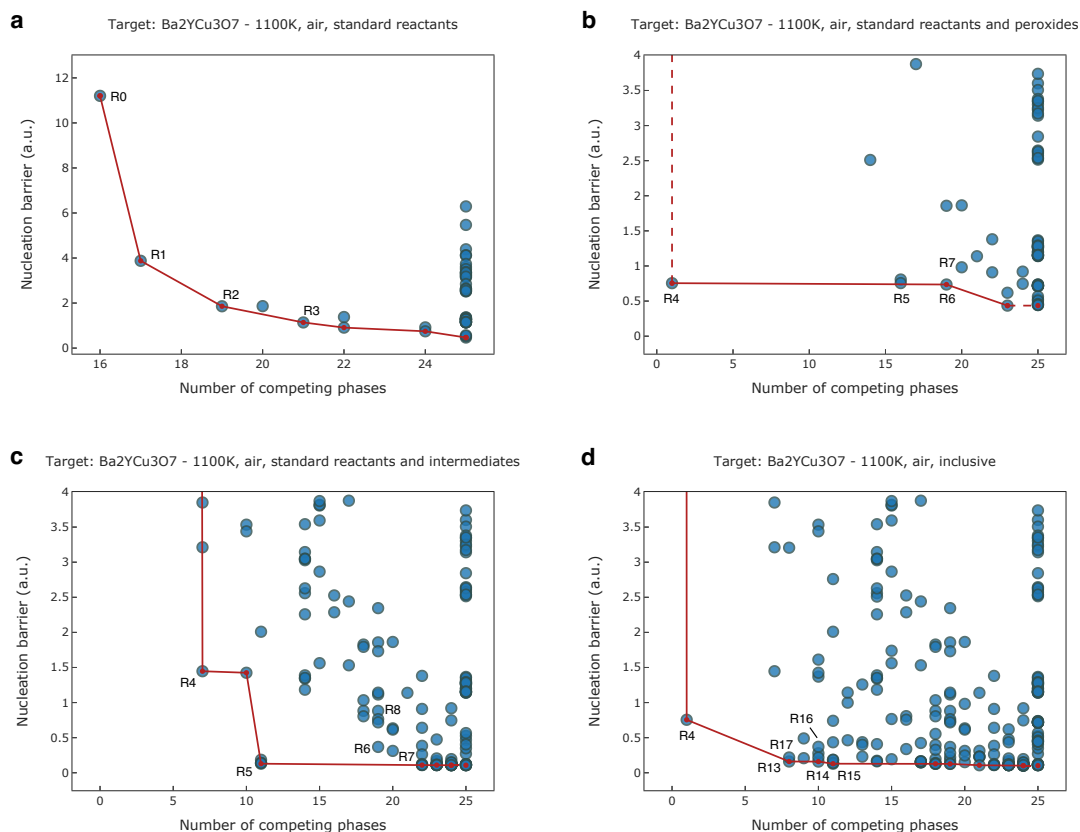


Figure 4: Route planning plots for synthesis of tetragonal YBa₂Cu₃O₇ under typical thermodynamic conditions using various sets of starting materials. Standard reactants include elements, stable binary compounds (excluding Y-Cu and Y-Ba alloys for simplicity) and carbonates. The set labeled as "inclusive" expands the set of standard reactants and peroxides to include ternary intermediates (e.g. Y₂Cu₂O₅). 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.

as metallic Cu,⁶⁹ Cu₂O^{70,71} and CuO^{72–75} 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 BaO₂-Cu₂O-Y₂O₃ route (R4), which was reported to be highly favorable for producing phase-pure YBCO with no carbonate impurities,^{70,71} 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 be 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,^{67,76–79} and can participate in reactions that lead to YBCO formation (e.g. R10 was reported previously⁶⁷). 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.

Case studies of other compounds

We show the route planning plots for several *less-common* inorganic compounds selected from the solid-state chemistry literature in Fig.5. These examples include Ca₃VN₃ by Vennos and DiSalvo,⁸⁰ Ca₂CrSbO₆ by Retuerto et al.,⁸¹ K₂V₃P₄O₁₇ by Lii et al.,⁸² LiNa₅Mo₉O₃₀ by Zhang et al.,⁸³ NaTi₈O₁₃ by Akimoto and Takei,⁸⁴ and Sr₂FeO₃F by Galasso and Darby⁸⁵. The selection was deliberately kept at a cursory level for broader testing, the only basic criteria being the reported compound to have a clear solid-state route in the cited study and a matching entry (calculation) in the MP database. In all examples, we find the reported routes to be on or close to the Pareto frontiers, providing further validation for the PIRO

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*	$3.0 \text{ CuO}(1692) + 0.25 \text{ O}_2 + 0.5 \text{ Y}_2\text{O}_3(2652) + 2.0 \text{ BaO}(1342)$
R1*	$1.0 \text{ O}_2 + 1.5 \text{ Cu}_2\text{O}(361) + 0.5 \text{ Y}_2\text{O}_3(2652) + 2.0 \text{ BaO}(1342)$
R2*	$3.0 \text{ Cu}(30) + 1.75 \text{ O}_2 + 0.5 \text{ Y}_2\text{O}_3(2652) + 2.0 \text{ BaO}(1342)$
R3*	$-2.0 \text{ CO}_2 + 3.0 \text{ CuO}(1692) + 0.25 \text{ O}_2 + 0.5 \text{ Y}_2\text{O}_3(2652) + 2.0 \text{ BaCO}_3(4559)$
R4*	$2.0 \text{ BaO}_2(1105) + 1.5 \text{ Cu}_2\text{O}(361) + 0.5 \text{ Y}_2\text{O}_3(2652)$
R5**	$3.0 \text{ CuO}(1692) + 0.5 \text{ Y}_2\text{O}_3(2652) + 0.5 \text{ BaO}_2(1105) + -1.5 \text{ CO}_2 + 1.5 \text{ BaCO}_3(4559)$
R6**	$1.5 \text{ Cu}(30) + 1.5 \text{ CuO}(1692) + 0.5 \text{ Y}_2\text{O}_3(2652) + 2.0 \text{ BaO}_2(1105)$
R7*	$2.0 \text{ BaO}_2(1105) + 3.0 \text{ CuO}(1692) + -0.75 \text{ O}_2 + 0.5 \text{ Y}_2\text{O}_3(2652)$
R8	$-0.5 \text{ CO}_2 + 0.6667 \text{ Ba}_2(\text{CuO}_2)_3(615789) + 0.1667 \text{ Ba}(122) + 0.5 \text{ BaCO}_3(4559) + 0.5 \text{ Y}_2\text{Cu}_2\text{O}_5(2882)$
R9	$2.25 \text{ O}_2 + -1.75 \text{ CO}_2 + 1.0 \text{ Ba}_2(\text{CuO}_2)_3(615789) + 0.25 \text{ Y}_4\text{C}_7(1200885)$
R10*	$0.25 \text{ O}_2 + 2.0 \text{ CuO}(1692) + -2.0 \text{ CO}_2 + 2.0 \text{ BaCO}_3(4559) + 0.5 \text{ Y}_2\text{Cu}_2\text{O}_5(2882)$
R11	$1.25 \text{ O}_2 + -2.0 \text{ CO}_2 + 2.0 \text{ Cu}(30) + 2.0 \text{ BaCO}_3(4559) + 0.5 \text{ Y}_2\text{Cu}_2\text{O}_5(2882)$
R12	$2.0 \text{ Cu}(30) + 1.25 \text{ O}_2 + 2.0 \text{ BaO}(1342) + 0.5 \text{ Y}_2\text{Cu}_2\text{O}_5(2882)$
R13	$1.0 \text{ Cu}(30) + 2.0 \text{ BaO}_2(1105) + 0.5 \text{ Cu}_2\text{O}(361) + 0.5 \text{ Y}_2\text{Cu}_2\text{O}_5(2882)$
R14	$1.5 \text{ Cu}(30) + 2.0 \text{ BaO}_2(1105) + 0.5 \text{ CuO}(1692) + 0.5 \text{ Y}_2\text{Cu}_2\text{O}_5(2882)$
R15	$1.75 \text{ BaO}_2(1105) + 0.25 \text{ BaY}_2\text{O}_4(3952) + 1.25 \text{ Cu}_2\text{O}(361) + 0.25 \text{ Y}_2\text{Cu}_2\text{O}_5(2882)$
R16	$1.0 \text{ BaO}_2(1105) + 0.5 \text{ Cu}_2\text{O}(361) + 0.5 \text{ Y}_2\text{Cu}_2\text{O}_5(2882) + 1.0 \text{ BaCuO}_2(997034)$
R17	$1.5 \text{ BaO}_2(1105) + 0.5 \text{ Cu}_2\text{O}(361) + 0.5 \text{ Ba}(\text{CuO})_2(7374) + 0.5 \text{ Y}_2\text{Cu}_2\text{O}_5(2882)$

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

approach presented.

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. 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.^{7,86} 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.

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 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).

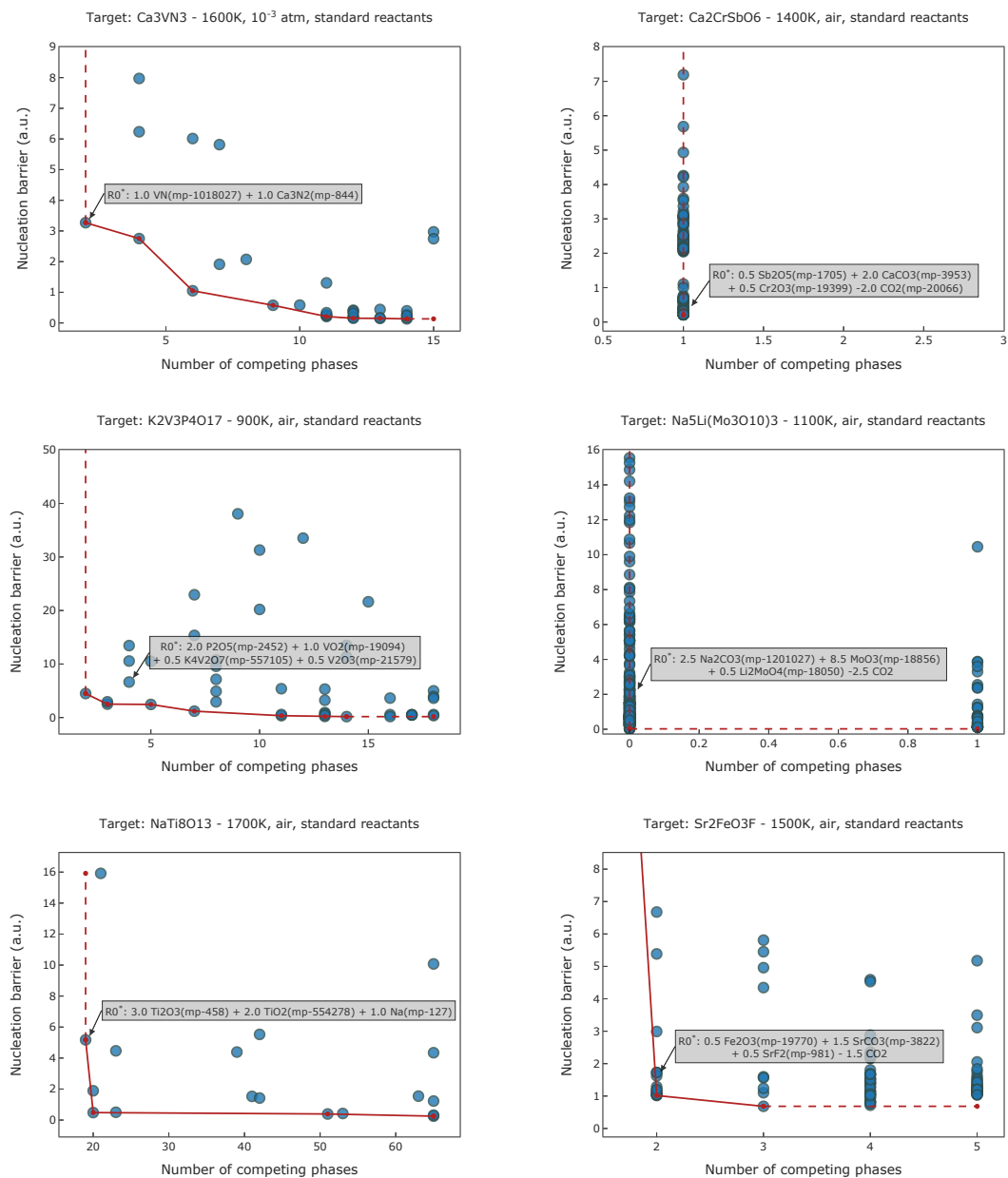


Figure 5: Route planning plots for solid-state synthesis various inorganic compounds from the chemistry literature. Routes reported in the literature are highlighted in the plots.

We exclude reactions that require making metal-alloys (e.g. YCu_2), except metal-carbides or elemental graphite to attain carbothermal⁸⁷ and/or self-sustaining high-temperature synthesis⁸⁸ 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.⁸⁹

The reaction tree obtained for solid-state synthesis of YBCO based on these guidelines is shown in Fig.6 and the respective reactions are listed listed in Table 4. In effect, 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 a ternary intermediate; including BaCuO_2 , Ba_2CuO_3 , $\text{Ba}(\text{CuO})_2$ and $\text{Y}_2\text{Cu}_2\text{O}_5$ except for the one-step peroxide route $\text{BaO}_2\text{-Y}_2\text{O}_3\text{-Cu}_2\text{O}$ (20674:R0). Many intermediate phases in fact evolve during synthesis of YBCO,^{67,77-79} and hence deliberately utilizing them can help constrain the reactions to follow a desired pathway. Besides, such ternary intermediates require two solid reactants in all high-ranking reactions in Table 4, and can therefore be produced at a single reaction interface more controllably. For instance, 20674:R1 uses

BaCuO₂, which is known to form easily during YBCO synthesis or explicitly from precursors CuO and BaCO₃,^{67,76} a high-ranking reaction in Table 4 for BaCuO₂ (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.⁶⁷ (only difference being use of CuO instead of 0.5 Cu₂O, which has negligible effect on reaction energy near temperatures close to CuO→Cu₂O decomposition). The BaCO₃ version of first-layer reaction 20674:R7 was also reported to yield YBCO.⁶⁷ The intermediate required for the first layer reaction 20674:R3 or 20674:R4; i.e., the Ba₂CuO₃ phase is reported to be amenable to synthesis via the peroxide route BaO₂-CuO (8790:R2) already ranked as favorable by the system,^{90,91} 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.

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

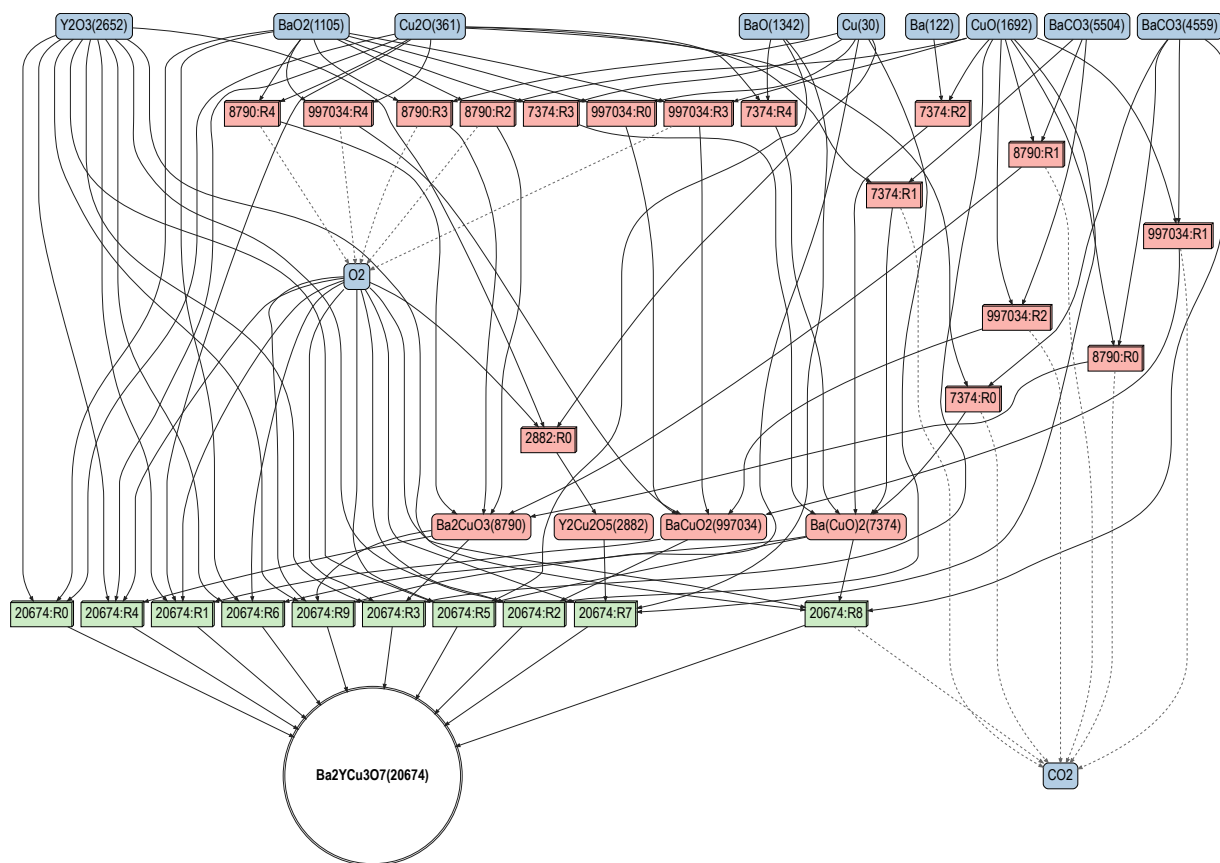


Figure 6: Retrosynthetic analysis for the planning of solid-state synthesis of YBCO with single-step or two-step pathways at 1100 K in open atmosphere, starting from simple precursors. Rounded rectangles show readily available precursor compounds/materials (blue rectangles) and intermediates (red rectangles) input for or output from synthesis reactions labeled in boxes. Reactions in red boxes produce intermediates, and those in green boxes produce the target compound. The particular form of a compound can be inferred from its MP entry number provided after its formula. Similarly, reaction labels contain the target compound, and the respective reaction number listed in the text. Gray dashed-lines indicate gas release. Equations of reactions are listed in Table 4.

Table 4: 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
<i>YBCO</i>	
20674:R0	0.5 Y ₂ O ₃ (2652) + 2.0 BaO ₂ (1105) + 1.5 Cu ₂ O(361)
20674:R1	0.5 O ₂ + 0.5 Y ₂ O ₃ (2652) + 0.5 Cu ₂ O(361) + 2.0 BaCuO ₂ (997034)
20674:R2	1.0 Cu(30) + 0.75 O ₂ + 0.5 Y ₂ O ₃ (2652) + 2.0 BaCuO ₂ (997034)
20674:R3	0.25 O ₂ + 2.0 CuO(1692) + 0.5 Y ₂ O ₃ (2652) + 1.0 Ba ₂ CuO ₃ (8790)
20674:R4	0.75 O ₂ + 0.5 Y ₂ O ₃ (2652) + 1.0 Cu ₂ O(361) + 1.0 Ba ₂ CuO ₃ (8790)
20674:R5	0.5 BaO(1342) + 1.5 Ba(CuO) ₂ (7374) + 1.0 O ₂ + 0.5 Y ₂ O ₃ (2652)
20674:R6	1.5 Ba(CuO) ₂ (7374) + 0.75 O ₂ + 0.5 Y ₂ O ₃ (2652) + 0.5 BaO ₂ (1105)
20674:R7	2.0 BaO(1342) + 0.25 O ₂ + 2.0 CuO(1692) + 0.5 Y ₂ Cu ₂ O ₅ (2882)
20674:R8	1.5 Ba(CuO) ₂ (7374) + 1.0 O ₂ + 0.5 Y ₂ O ₃ (2652) + -0.5 CO ₂ + 0.5 BaCO ₃ (4559)
20674:R9	2.0 Cu(30) + 1.25 O ₂ + 0.5 Y ₂ O ₃ (2652) + 1.0 Ba ₂ CuO ₃ (8790)
<i>Intermediates</i>	
Ba(CuO) ₂	
7374:R0	-1.0 CO ₂ + 1.0 Cu ₂ O(361) + 1.0 BaCO ₃ (4559)
7374:R1	1.0 BaCO ₃ (5504) + -1.0 CO ₂ + 1.0 Cu ₂ O(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 ₂ O(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 ₂ O(361) + -0.25 O ₂
Ba ₂ CuO 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 ₂ O(361) + -0.75 O ₂ + 2.0 BaO ₂ (1105)
Y ₂ Cu ₂ O ₅	
2882:R0	1.0 Y ₂ O ₃ (2652) + 1.0 O ₂ + 2.0 Cu(30)

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. 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.⁹² DFT energies for solids can be improved through use of new functionals in high-throughput datasets.⁹³ 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.⁹⁴

As noted by Jansen and Schön two decades ago,^{19,95} 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 in Fig. 5) and pave the way for more effective planning of synthesis. Today, this mapping can be achieved by efficiently populating the space with low-energy hypothetical compounds generated using data-driven methods.^{17,96–98} 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.^{99–102}

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, 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 com-

pounds. 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.¹⁰³

References

- (1) Kanatzidis, M. G. et al. *Progress in Solid State Chemistry* **2008**, *36*, 1–133.
- (2) Kohlmann, H. *European Journal of Inorganic Chemistry* **2019**, *2019*, 4174–4180.
- (3) Chamorro, J. R.; McQueen, T. M. *Accounts of Chemical Research* **2018**, *51*, 2918–2925.
- (4) Martinolich, A. J.; Neilson, J. R. *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. *The Journal of Chemical Physics* **1949**, *17*, 71–73.
- (7) Corey, E. J. *Angewandte Chemie International Edition in English* **1991**, *30*, 455–465.
- (8) Kim, E.; Huang, K.; Saunders, A.; McCallum, A.; Ceder, G.; Olivetti, E. *Chemistry of Materials* **2017**, *29*, 9436–9444.
- (9) Kim, E.; Huang, K.; Kononova, O.; Ceder, G.; Olivetti, E. *Matter* **2019**, *1*, 8–12.

- (10) Kononova, O.; Huo, H.; He, T.; Rong, Z.; Botari, T.; Sun, W.; Tshitoyan, V.; Ceder, G. *Scientific data* **2019**, *6*, 203.
- (11) Oliynyk, A. O.; Antono, E.; Sparks, T. D.; Ghadbeigi, L.; Gaultois, M. W.; Meredig, B.; Mar, A. *Chemistry of Materials* **2016**, *28*, 7324–7331.
- (12) Aykol, M.; Hegde, V. I.; Hung, L.; Suram, S.; Herring, P.; Wolverton, C.; Hummelshøj, J. S. *Nature Communications* **2019**, *10*, 2018.
- (13) Malik, S. A.; Goodall, R. E. A.; Lee, A. A. **2020**, arXiv 2007.15752.
- (14) Jang, J.; Gu, G. H.; Noh, J.; Kim, J.; Jung, Y. *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. *Science Advances* **2016**, *2*, e1600225–e1600225.
- (16) Aykol, M.; Dwaraknath, S. S.; Sun, W.; Persson, K. A. *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. *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. *Nature Materials* **2020**,
- (19) Jansen, M. *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. *Journal of Physical Chemistry B* **2006**, *110*, 17315–17328.
- (23) Turnbull, D.; Vonnegut, B. *Industrial & Engineering Chemistry* **1952**, *44*, 1292–1298.
- (24) Becker, R.; Döring, W. *Annalen der Physik* **1935**, *416*, 719–752.
- (25) Russell, K. C. *Advances in Colloid and Interface Science* **1980**, *13*, 205–318.
- (26) Kalikmanov, V. *Nucleation Theory*; 2013.
- (27) Boldyrev, V. V. *Reactivity of Solids* **1990**, *8*, 231–246.
- (28) Merkle, R.; Maier, J. *Zeitschrift für 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. *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. *J. Phys. Chem. Ref. Data* **1982**, *Vol. 11*, Suppl. 2.
- (32) Wang, L.; Maxisch, T.; Ceder, G. *Phys. Rev. B* **2006**, *73*, 195107.
- (33) Grindy, S.; Meredig, B.; Kirklin, S.; Saal, J.; Wolverton, C. *Physical Review B* **2013**, *87*, 075150.
- (34) Zur, A.; McGill, T. C. *Journal of Applied Physics* **1984**, *55*, 378–386.
- (35) Ding, H.; Dwaraknath, S. S.; Garten, L.; Ndione, P.; Ginley, D.; Persson, K. A. *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. *Computational Materials Science* **2013**, *68*, 314–319.
- (37) Ward, L.; Liu, R.; Krishna, A.; Hegde, V. I.; Agrawal, A.; Choudhary, A.; Wolverton, C. *Physical Review B* **2017**, *96*, 024104.
- (38) Ward, L. et al. *Computational Materials Science* **2018**, *152*, 60–69.
- (39) Belsky, A.; Hellenbrandt, M.; Karen, V. L.; Luksch, P. *Acta Crystallographica Section B Structural Science* **2002**, *58*, 364–369.
- (40) Vijatović, M. M.; Bobić, J. D.; Stojanović, B. D. *Science of Sintering* **2008**, *40*, 155–165.
- (41) Stojanovic, B. D.; Simoes, A. Z.; Paiva-Santos, C. O.; Jovalekic, C.; Mitic, V. V.; Varela, J. A. *Journal of the European Ceramic Society* **2005**, *25*, 1985–1989.
- (42) Cournil, M.; Soustelle, M.; Thomas, G. *Oxidation of Metals* **1979**, *13*, 89–104.
- (43) Buscaglia, M. T.; Bassoli, M.; Buscaglia, V.; Alessio, R. *Journal of the American Ceramic Society* **2005**, *88*, 2374–2379.
- (44) Beauger, A.; Mutin, J. C.; Niepce, J. C. *Journal of Materials Science* **1983**, *18*, 3543–3550.
- (45) Lee, T. T.; Huang, C. Y.; Chang, C. Y.; Lin, S. P.; Su, C. Y.; Lee, C. T.; Fujimoto, M. *Japanese Journal of Applied Physics* **2011**, *50*.
- (46) Komarov, A. V.; Parkin, I. P. *Polyhedron* **1996**, *15*, 1349–1353.
- (47) Licheri, R.; Fadda, S.; Orrù, R.; Cao, G.; Buscaglia, V. *Journal of the European Ceramic Society* **2007**, *27*, 2245–2253.

- (48) Larson, E. M.; Wong, J.; Holt, J. B.; Waide, P. A.; Rupp, B. *Powder Diffraction* **1999**, *14*, 111–113.
- (49) Lotnyk, A.; Senz, S.; Hesse, D. *Solid State Ionics* **2006**, *177*, 429–436.
- (50) Rössel, M.; Gablenz, S.; Müller, T.; Röder, A.; Abicht, H. P. *Analytical and Bioanalytical Chemistry* **2003**, *375*, 310–314.
- (51) Rössel, M.; Höche, H. R.; Leipner, H. S.; Völtzke, D.; Abicht, H. P.; Hollricher, O.; Müller, J.; Gablenz, S. *Analytical and Bioanalytical Chemistry* **2004**, *380*, 157–162.
- (52) Beauger, A.; Mutin, J. C.; Niepce, J. C. *Journal of Materials Science* **1983**, *18*, 3041–3046.
- (53) Pfaff, G. *Journal of Materials Science Letters* **1991**, *10*, 129–131.
- (54) Zhu, N.; West, A. R. *Journal of the American Ceramic Society* **2010**, *93*, 295–300.
- (55) MIZUSHIMA, K.; JONES, P.; WISEMAN, P.; GOODENOUGH, J. *Solid State Ionics* **1981**, *3-4*, 171–174.
- (56) Antolini, E. *Journal of the European Ceramic Society* **1998**, *18*, 1405–1411.
- (57) Antolini, E. *Solid State Ionics* **2004**, *170*, 159–171.
- (58) Lundblad, A.; Bergman, B. *Solid State Ionics* **1997**, *96*, 173–181.
- (59) Carewska, M.; Di Bartolomeo, A.; Scaccia, S. *Thermochimica Acta* **1995**, *269-270*, 491–506.
- (60) Berbenni, V.; Milanese, C.; Bruni, G.; Marini, A. *Materials Chemistry and Physics* **2006**, *100*, 251–256.
- (61) Johnston, W. D.; Heikes, R. R.; Sestrich, D. *Journal of Physics and Chemistry of Solids* **1958**, *7*, 1–13.

- (62) Kobayashi, H.; Tsukasaki, T.; Ogasawara, Y.; Hibino, M.; Kudo, T.; Mizuno, N.; Honma, I.; Yamaguchi, K. *ACS Applied Materials and Interfaces* **2020**, *12*, 43605–43613.
- (63) Möller, A. *Chemistry of Materials* **1998**, *10*, 3196–3201.
- (64) Amatucci, G.; Tarascon, J.; Klein, L. *J. Electrochem. Soc.* **1996**, *143*, 1114–1123.
- (65) 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. *Physical Review Letters* **1987**, *58*, 908–910.
- (66) Hrovat, M.; Bernik, S.; Kolar, D.; Krasevec, V. *International Journal of Modern Physics B* **1987**, *01*, 1027–1033.
- (67) Ruckenstein, E.; Narain, S.; Wu, N.-L. *Journal of Materials Research* **1989**, *4*, 267–272.
- (68) Pathak, L. C.; Mishra, S. K. *Superconductor Science and Technology* **2005**, *18*.
- (69) Lebrat, J. P.; Varma, A. *Combustion Science and Technology* **1993**, *88*, 177–185.
- (70) Kao, M. *Materials Letters* **1987**, *6*, 53–57.
- (71) Kao, M.; McKinney, B. *Materials Letters* **1991**, *11*, 91–95.
- (72) Hepp, A.; Gaier, J. *Materials Research Bulletin* **1988**, *23*, 693–700.
- (73) Costa, C. A.; Ferretti, M.; Olcese, C. L.; Cimberle, M. R.; Ferdeghini, C.; Nicchiotti, G. L.; Siri, A. S.; Rizzuto, C. *Journal of Crystal Growth* **1987**, *85*, 623–627.
- (74) Fahlman, B. D. *Journal of Chemical Education* **2001**, *78*, 1182.
- (75) 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. **2020**,

- (76) Flor, G.; Scavini, M.; Anselmi-Tamburini, U.; Spinolo, G. *Solid State Ionics* **1990**, *43*, 77–83.
- (77) Peterson, D. E.; Kubat-Martin, K. A.; George, T. G.; Zocco, T. G.; Thompson, J. D. *Journal of Materials Research* **1991**, *6*, 11–17.
- (78) Cogdell, C.; Wayment, D.; Kubat-Martin, K. *J. Chem. Educ.* **1995**, *72*, 840–841.
- (79) Anselmi-Tamburini, U.; Ghigna, P.; Spinolo, G.; Flor, G. *Journal of Physics and Chemistry of Solids* **1991**, *52*, 715–721.
- (80) Vennos, D. A.; DiSalvo, F. J. *Journal of Solid State Chemistry* **1992**, *98*, 318–322.
- (81) Retuerto, M.; Alonso, J. A.; García-Hernández, M.; Martínez-Lope, M. J. *Solid State Communications* **2006**, *139*, 19–22.
- (82) Lii, K. H.; Tsai, H. J. *Journal of Solid State Chemistry* **1990**, *401*, 396–401.
- (83) Zhang, W.; Yu, H.; Cantwell, J.; Wu, H.; Poeppelmeier, K. R.; Halasyamani, P. S. *Chemistry of Materials* **2016**, *28*, 4483–4491.
- (84) Akimoto, J.; Takei, H. *Journal of Solid State Chemistry* **1991**, *90*, 147–154.
- (85) Galasso, F.; Darby, W. *Journal of Physical Chemistry* **1963**, *67*, 1451–1453.
- (86) Coley, C. W.; Rogers, L.; Green, W. H.; Jensen, K. F. *ACS Central Science* **2017**, *3*, 1237–1245.
- (87) Martirosyan, K. S.; Luss, D. *AIChE Journal* **2005**, *51*, 2801–2810.
- (88) Munir, Z. A.; Anselmi-Tamburini, U. *Materials Science Reports* **1989**, *3*, 279–365.
- (89) Hwang, C.-L.; Lai, Y.-J.; Liu, T.-Y. *Computers & Operations Research* **1993**, *20*, 889–899.

- (90) Armstrong, A. R.; Janes, R.; Singh, K. K.; Edwards, P. P. *Bull. Mater. Sci.* **1991**, *14*, 641–649.
- (91) Umezawa, A.; Zhang, W.; Gurevich, A.; Feng, Y.; Hellstrom, E. E.; Larbalestier, D. C. *Nature* **1993**, *364*, 129–131.
- (92) Bartel, C. J.; Millican, S. L.; Deml, A. M.; Rumpitz, J. R.; Tumas, W.; Weimer, A. W.; Lany, S.; Stevanović, V.; Musgrave, C. B.; Holder, A. M. *Nature Communications* **2018**, *9*.
- (93) Zhang, Y.; Kitchaev, D. A.; Yang, J.; Chen, T.; Dacek, S. T.; Sarmiento-Pérez, R. A.; Marques, M. A.; Peng, H.; Ceder, G.; Perdew, J. P.; Sun, J. *npj Computational Materials* **2018**, *4*.
- (94) McDermott, M.; Dwaraknath, S.; Persson, K. **2020**,
- (95) Christian Schön, J.; Jansen, M. *Angewandte Chemie (International Edition in English)* **1996**, *35*, 1286–1304.
- (96) Woodley, S. M.; Catlow, R. *Nature materials* **2008**, *7*, 937–46.
- (97) Hautier, G.; Fischer, C. C.; Jain, A.; Mueller, T.; Ceder, G. *Chemistry of Materials* **2010**, *22*, 3762–3767.
- (98) Montoya, J. H.; Winther, K. T.; Flores, R. A.; Bligaard, T.; Hummelshøj, J. S.; Aykol, M. *Chemical Science* **2020**, 8517–8532.
- (99) 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. *Computational Materials Science* **2012**, *58*, 227–235.
- (100) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. *npj Computational Materials* **2015**, *1*, 15010.

- (101) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. *JOM* **2013**, *65*, 1501–1509.
- (102) Choudhary, K. et al. *npj Computational Materials* **2020**, *6*, 173.
- (103) Momma, K.; Izumi, F. *Journal of Applied Crystallography* **2011**, *44*, 1272–1276.