Assessing Thermodynamic Selectivity of Solid-State Reactions for the Predictive Synthesis of Inorganic Materials

Matthew J. McDermott,[†,‡,@] Brennan C. McBride,[¶,@] Corlyn Regier,[¶] Gia Thinh Tran,[¶] Yu Chen,[†,‡] Adam A. Corrao,[§] Max C. Gallant,[†,‡] Gabrielle E. Kamm,[§] Christopher J. Bartel,[∥] Karena W. Chapman,[§] Peter G. Khalifah,[§,⊥] Gerbrand Ceder,[†,‡] James R. Neilson,[¶] and Kristin A. Persson[*,#,‡]

†Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA
‡Department of Materials Science and Engineering, University of California, Berkeley, CA
¶Department of Chemistry, Colorado State University, Fort Collins, CO
§Department of Chemistry, Stony Brook University, Stony Brook, NY
∥Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN
⊥Chemistry Division, Brookhaven National Laboratory, Upton, NY
#Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA
@These authors contributed equally to this work.

E-mail: kapersson@lbl.gov

Abstract

Synthesis is a major challenge in the discovery of new inorganic materials. There is currently limited theoretical rationale for planning optimal solid-state synthesis procedures that selectively yield desired targets with minimal impurities. Using an interface reaction model, we propose two selectivity metrics – primary and secondary competition – to measure the favorability of target and impurity phase formation in solid-state reactions. We first apply these metrics to assess the selectivity of 3,520 solid-state synthesis recipes extracted from the scientific literature, comparing the thermodynamic optimality of various approaches to popular targets (e.g., $LiMn_2O_4$, $LiFePO_4$) under our framework. Secondly, we implement these metrics in a data-driven synthesis planning workflow and test its use in suggesting synthetic routes to barium titanate ($BaTiO_3$). From an 18-element chemical reaction network created with first-principles materials thermodynamic data from the Materials Project, we identify 82,985 synthesis reactions for $BaTiO_3$ and select nine to test experimentally, characterizing their reaction pathways using *ex post facto* synchrotron powder X-ray diffraction over a wide range of temperatures. Analysis of experimental results indicates that our proposed selectivity metrics correlate with observed target and impurity formation. We identify two favorably predicted reactions with unconventional $BaS/BaCl_2$ and Na_2TiO_3 precursors that yield $BaTiO_3$ more rapidly and with fewer impurities than conventional solid-state approaches, highlighting the importance of precursor selection and consideration of complex chemistries with additional elements beyond those in the target material. Altogether, our framework serves as a foundation for the predictive synthesis of inorganic materials, facilitating the optimization of existing synthesis approaches and the design of syntheses to novel materials, including those that cannot be easily made from conventional "off-the-shelf" precursors.

Introduction

The predictive synthesis of inorganic materials remains a grand challenge in modern chemistry and materials science.[1] Unlike organic synthesis, which is often described via discrete reaction steps or mechanisms, inorganic materials synthesis reactions cannot be deconstructed into elementary steps,[2,3] hindering the analogous development of retrosynthetic analysis techniques[4] and computer-aided synthesis planning.[5] This lack of successful mechanistic models has made the synthesis of predicted new materials a critical bottleneck in high-throughput computational materials design efforts,[6] with many proposed materials having yet to be successfully synthesized.[7–9]

While there are numerous inorganic synthesis methods (e.g., hydrothermal, mechanochemical, sol-gel, intercalation, etc.),[10] we limit the scope of this work to bulk solid-state synthesis via powder reactions. This choice has been motivated by the straightforward and scalable nature of working with bulk powders, which makes solid-state synthesis suitable for applications ranging from one-off laboratory synthesis to industrial mass manufacturing. In powder reactions, product formation proceeds via nucleation and growth at interfacial contact areas in the powder mixture (Figure 1a).[11] The equilibrium phases of the reacting system can be predicted by constructing a convex hull in free energy and composition space, where the composition axis is a mixing ratio between the two precursor compositions (Figure 1b).[12] Here, we calculate the convex hull exclusively using normalized compositions and energies (i.e., on a per-atom basis). This construction, which we refer to herein as the "interface reaction hull", is a subsection of the compositional phase diagram for binary systems and a "quasibinary" two-dimensional slice of the full phase diagram for chemical systems with three or more elements. The exact product species, and the sequence in which they appear, cannot be predicted with thermodynamics alone; to do so requires intimate knowledge of the kinetic rates of all physically feasible reactions. However, a commonly adopted theoretical simplification assumes that the reaction product(s) with the most negative *pairwise* reaction energy will be the first to nucleate and grow as a powder mixture is heated.[11,13] This

hypothesis is based on two principles: 1) the random packing of solid crystallites results in very few locations where three or more particles are simultaneously in contact, and 2) the activation energy barrier to nucleation scales inversely with free energy as $\Delta G^\dagger \propto \frac{\gamma^3}{\Delta G_{rxn}^2}$. The surface energy, γ, is particularly important when comparing the feasibility of reactions with similar energies.[14] Despite this, many solid-state reactions are likely not nucleation-limited but rather *transport-limited* due to the relative sluggishness of solid-state diffusion and the often large driving forces of these reactions (10-100 kJ/mol).[15]
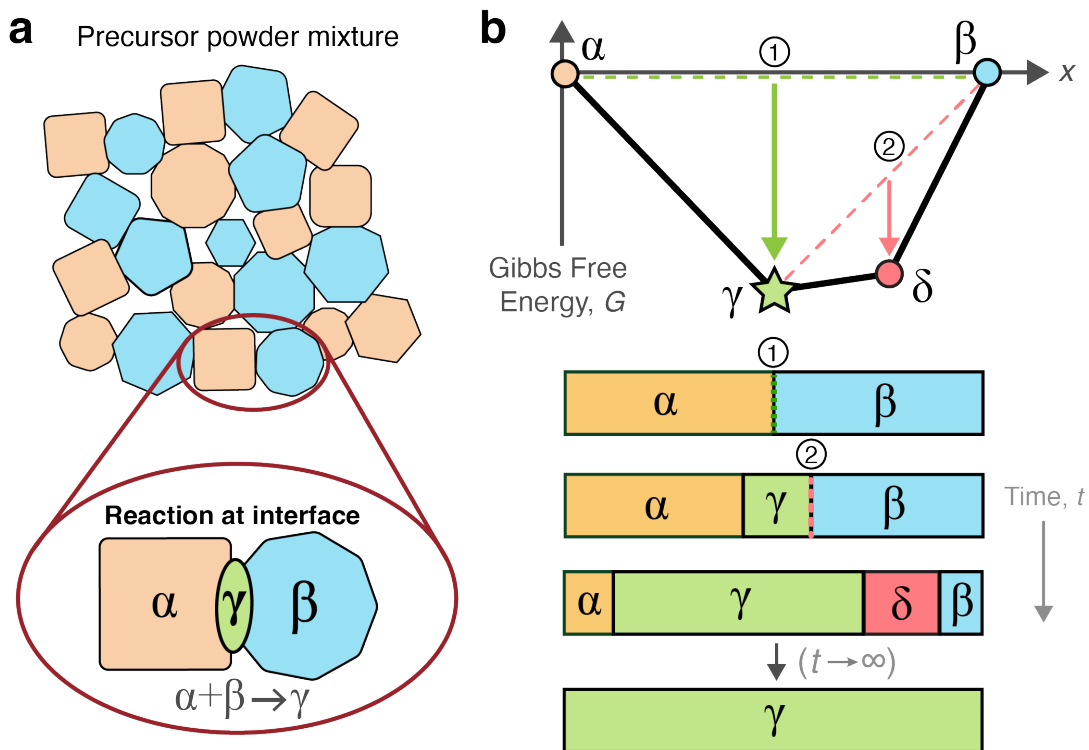
Figure 1: Modeling chemical reactions at heterogeneous solid interfaces in a binary/quasibinary chemical system. (a) Cartoon model of a powder reaction between the hypothetical precursors: α (orange) and β (blue). The nucleation of a new phase, γ (green), occurs at the $\alpha|\beta$ interface according to the reaction $\alpha + \beta \longrightarrow \gamma$. (b) A possible reaction pathway for the powder system in which a secondary reaction of the equilibrium phase (γ) yields an impurity phase, δ (red). The interface reaction hull (top) shows available interfacial reactions and their corresponding Gibbs free energies, G, and mixing ratios, x. The one-dimensional spatial model (bottom) shows reaction steps beginning from an equal mixture of α and β. The impurity phase, δ, may be kinetically retained in a local equilibrium state; however, with infinite time, the system should approach the global equilibrium state composed entirely of γ.

The complex interplay between thermodynamics and kinetics makes solid-state synthesis prone to the unpredictable and undesirable formation of impurity phases.[13] A classic example of a nonselective synthesis is that of the prototypical multiferroic bismuth ferrite, $BiFeO_3$, via the standard reaction from binary oxides: $Bi_2O_3 + Fe_2O_3 \longrightarrow 2\ BiFeO_3$. This reaction typically yields impurity phases $Bi_2Fe_4O_9$ and $Bi_{25}FeO_{39}$, which are challenging to isolate and remove.[16,17] Unfortunately, the presence of an impurity phase is difficult to predict *a priori*, and is typically attributed to "kinetic" factors or changes in phase equilibria related to precursor purity, morphology, volatility, or processing conditions. To optimize the performance of solid-state reactions and maximize conversion to the desired target, the experimentalist frequently relies on intuition and heuristic rules to choose the 1) precursor compositions (typically off-the-shelf binary phases such as carbonates, oxides, etc.), 2) grinding/milling protocol, 3) synthesis annealing temperature, 4) synthesis atmosphere (e.g., vacuum, flowing O_2), 5) synthesis time, and 6) cooling protocol. Heuristics include well-known rules such as Tamman's Rule for estimating reaction onset temperature (i.e., two-thirds the melting temperature of the precursor with the lowest melting point),[18] as well as "chemical intuition" or human-biased experimental protocols (e.g., selecting synthesis times based on common increments, such as four hours, eight hours, etc.).[19] Unfortunately, these heuristics may be insufficient to achieve successful synthesis on the first attempt(s), necessitating follow-up experiments that can be time-intensive and costly. In the worst cases, human-biased heuristics lead to lower success rates than randomly-generated experimental protocols.[20]

The *a priori* calculation of reaction selectivity in solid-state synthesis permits the ranking of synthesis approaches based on their thermodynamic likelihood of success, thereby circumventing the current time-consuming trial-and-error (Edisonian) approach. The assessment of reaction selectivity is particularly relevant in the proposal of optimal synthesis precursors;[21] in several cases, improved navigation of the phase diagram was shown to lead to a more practical synthesis.[11,22–25] However, no solid-state reaction selectivity metric has been formally established. In recent work,[14] Aykol et al. demonstrated a computational workflow for

ranking solid-state synthesis reactions by two performance metrics: 1) a catalytic nucleation barrier factor incorporating structural similarity and epitaxial matching, and 2) the number of known competing phases. These metrics perform well in rationalizing successful syntheses in the literature but lack generality; for example, the nucleation metric is derived assuming all reactions are nucleation-limited, which, as discussed, is not true for many solid-state reactions. Additionally, while a metric based on the total *number* of competing phases is significant as it hints at a measure of reaction selectivity, such a scheme does not account for the *relative* stability of these competing phases. A count-based selectivity metric is also biased by how many phases are known to exist at the present time and the extent to which various structural configurations (e.g., disordered or defective phases) have been enumerated within the data.

In this work, we address the longstanding issue of assessing the selectivity of solid-state reactions by deriving two complementary thermodynamic metrics measuring the degree of phase competition from the interface reaction model. We incorporate these competition metrics into a computational synthesis planning workflow for identifying and ranking synthesis reactions, which builds upon the high-throughput reaction enumeration tools we previously developed for constructing solid-state chemical reaction networks[26] from large materials databases such as the Materials Project.[27] Our selectivity metrics, computational workflow, literature analysis, and experimental findings yield a framework for generating more optimal and efficient solid-state synthesis routes, providing a foundation for the predictive synthesis of inorganic materials. The suggestion of nonstandard precursors, particularly those involving additional elements beyond those in the target composition, expands the synthetic capabilities of the solid-state approach.

Results and Discussion

Derivation of Selectivity Metrics

The Interface Reaction Hull

To construct the interface reaction hull (Figure 1b), one begins with thermodynamic data for the reacting system (i.e., a set of relevant phases and their compositions and energies). In this work, we acquire formation enthalpies, ΔH_f, from the Materials Project database[27] and extend them to Gibbs free energies of formation, $\Delta G_f(T)$, through the use of a prior machine learning model[28] and supplemental experimental thermochemistry data[29] (see Methods). For systems with two elements, the interface reaction hull is equivalent to the binary compositional phase diagram, where each vertex represents a single phase. However, for systems with three or more elements, the non-precursor vertices include both single phases and mixtures of phases. These mixtures are stoichiometric combinations of phases representing the products of balanced reactions of the precursors. The balanced reactions can be determined via 1) computing slices of the full compositional phase diagram along the tie-line connecting the precursors[12] or 2) combinatorial reaction enumeration.[26] More specifically, the maximum number of products for a particular vertex is one less than the number of elements in the system. The interface reaction hull is thus generalized such that all non-precursor vertices correspond to reactions with coordinates given by the atomic mixing ratio of precursors, x, and the Gibbs free energy of the reaction, ΔG_{rxn}. This model can be further generalized to environmental conditions other than fixed temperature and pressure by constructing the hull with the appropriate thermodynamic potential. For example, in open systems, one would use the grand potential energy, Φ. Note that in these systems, the hull vertices may include additional open elemental reactants/products (e.g., O_2) that do not factor into the determination of x.

As a model for solid-state reactions, the interface reaction hull construction also rationalizes the formation of impurity phases. To demonstrate this, we revisit the binary system

7

in Figure 1. In this system, the target phase (γ) is predicted to form first because it is the phase with the highest driving force of formation (most negative ΔG_{rxn}), irrespective of the average composition of the total system.[12] When the γ phase forms, however, it also introduces two additional interfaces: $\alpha|\gamma$ and $\gamma|\beta$. These secondary interfaces can be modeled via the construction of new interface reaction hulls or, more simply, by splitting the original hull into two subsections (i.e., to the left and the right of the target). Figure 1b suggests that the $\gamma|\beta$ interface should produce an impurity phase via the exergonic reaction, $\gamma + \beta \longrightarrow \delta$. The full conversion of reactants to the target phase is impeded while δ persists. In a "one-dimensional" solid-state reaction (Figure 1b, bottom), local thermodynamic equilibrium may be achieved when the system reaches a state in which all stable product phases on the interface reaction hull have formed and the growth of the product layer(s) slows down until it ceases entirely. This situation has been observed in previous experimental studies on diffusion couples.[30,31] This observed mixture of products may be kinetically "stable" (i.e., with locally stable interfaces), but it is not the global equilibrium state of the system. Rather, the global equilibrium state is the combination of phases that minimizes the free energy given the composition of the entire powder mixture; in Figure 1b, this corresponds to entirely γ. In powder reactions, access to the global equilibrium state is often provided by re-grinding steps in which new interfaces are exposed and mixed to facilitate complete conversion to the equilibrium products. However, in reacting systems with significant phase competition and slow transport, high temperatures and long heating times may be necessary but impractical; pure target synthesis may not be achievable if the desired products are unstable at high temperatures. These situations can be avoided entirely by proposing alternative precursors that are more selective (i.e., those with interface reaction hulls containing few to no competing phases).

Measuring Phase Competition

To predict the thermodynamic selectivity of a solid-state reaction, we propose two complementary metrics for assessing phase competition using the interface reaction hull: primary (C_1) and secondary (C_2) competition. Although both metrics measure the relative energetic favorability of competing reactions, they model different mechanisms for impurity formation. The primary competition measures the favorability of competing reactions of the original precursors, while the secondary competition measures the favorability of *subsequent* competing reactions between the precursors and target phase(s). The origin of the two competition mechanisms is illustrated in Figure 2.
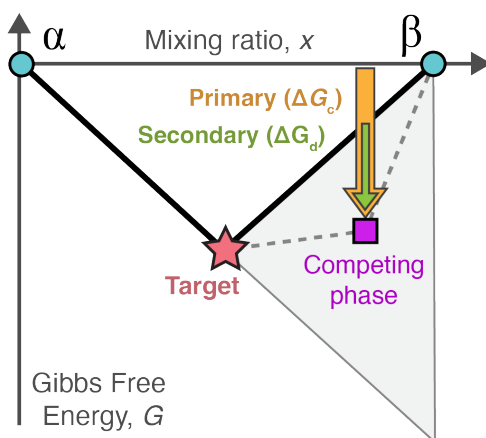
Figure 2: Origin of primary and secondary competition in solid-state reactions. In this simple interface reaction hull for a binary (two-element) system, two interface reactions can form a competing phase (magenta square). The primary reaction (yellow arrow) occurs at the interface between the two precursors α and β, whereas the secondary reaction (green arrow) occurs between the target phase (pink star) and the remaining β precursor, leading to a smaller driving force (arrow length). The coordinates of the competing phase, which must lie within the illustrated bounds (gray triangle) if the target phase is to be stable, determine the relative values of the primary and secondary reaction energies.

Primary competition, C_1, is measured via calculation of the relative thermodynamic advantage of the most exergonic competing reaction from the original precursors, as assessed

through an energy difference:

$$C_1 = \Delta G_{\mathrm{rxn}} - \min_i(\Delta G_{c_i}). \tag{1}$$

Here, ΔG_{rxn} is the energy of the target synthesis reaction, and ΔG_{c_i} are the energies of possible competing reactions from the precursors. Lower C_1 values are favorable and result in more selective target formation. When C_1 is positive, the target reaction is less energetically favorable than the competing reaction with the greatest driving force (most negative energy), suggesting that a competing phase is likely to form. On the other hand, when C_1 is negative, the target reaction is predicted to have the greatest driving force of any reaction on the hull. By considering only the single most competitive reaction, this functional form avoids the aforementioned bias related to using the total *number* of competing reactions. When no exergonic competing reactions are predicted for an interface, the competing reaction energy term is assigned a value of zero, representing the scenario in which the precursors do not react (e.g., $\alpha \longrightarrow \alpha$). This results in the limiting condition: $C_1 \geq \Delta G_{\mathrm{rxn}}$.

Secondary competition, C_2, assesses the favorability of impurity phase formation via secondary reactions between the target and precursor(s). This metric is important and distinct from primary competition because it measures the relative stability of the products of the target synthesis reaction with respect to decomposition into the competing phase(s). Their relative stability can be measured by computing the "inverse distance to the hull" (Figure 2), which for systems with one competing phase, is equivalent to the secondary reaction energy, ΔG_d, at the precursor-target interface.

In an interface reaction hull with several competing phases, a sequence of multiple secondary reactions may occur (Figure 3). When the target phase is formed, it introduces two new precursor-target interfaces that divide the hull into subsections to the left and right of the target. A secondary reaction may occur in either subsection, exposing another two interfaces. If, at either interface, there is a remaining driving force to form additional competing
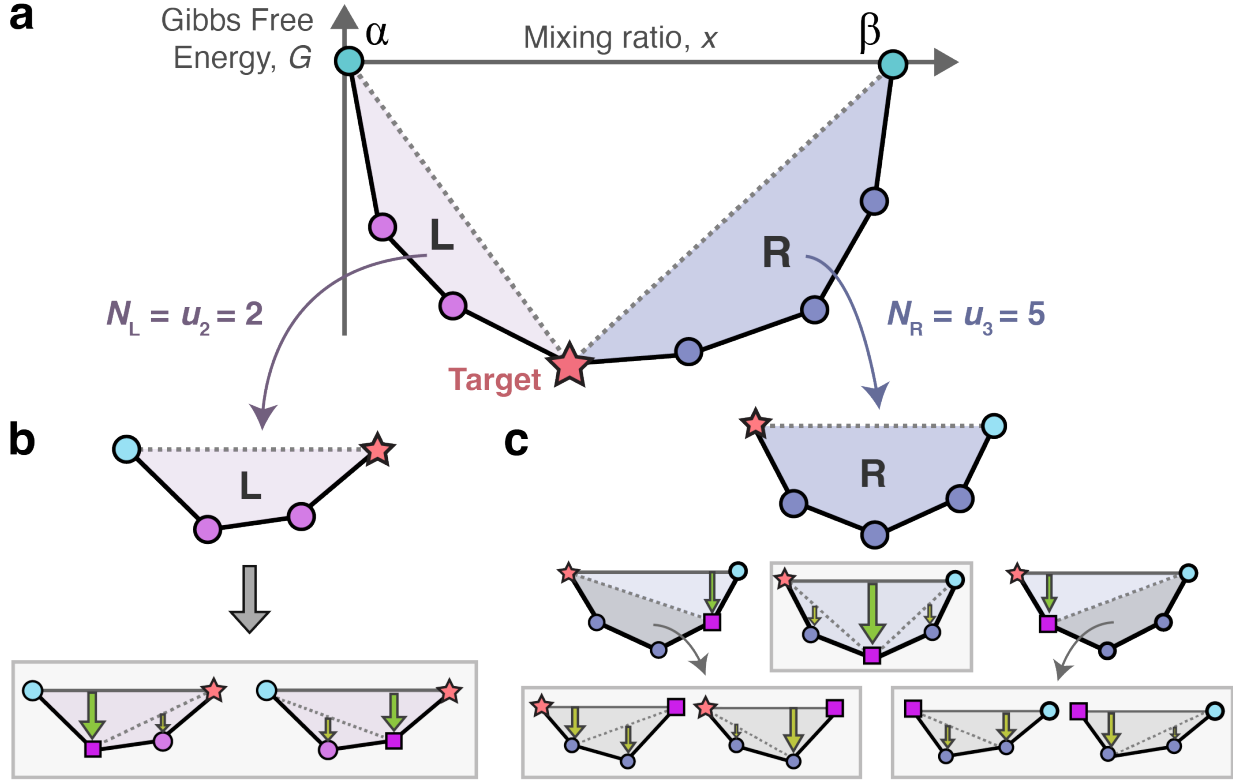
Figure 3: Secondary reaction sequences in an interface reaction hull. (a) The hull is divided into two subsections to the left (L) and right (R) of the target, representing the two additional precursor-target interfaces. (b-c) Secondary reaction sequences on each side of the target, with gray boxes highlighting the final reaction sequences. The recursive, binary nature results in the number of unique sequences, $N(n)$, following the Catalan numbers u_n.

phases, then this process may continue in a recursive fashion until all possible secondary reactions have occurred. There are multiple ways to draw a feasible secondary reaction sequence (Figures 3b,c). Consider a particular secondary reaction sequence indexed j. This sequence has a total energy given by the sum of the energies of its n steps,

$$\Delta G_{2,j} = \sum_{k=1}^{n} \Delta G_{d_{j,k}} = \Delta G_{d_{j,1}} + \Delta G_{d_{j,2}} + ... + \Delta G_{d_{j,n}}, \tag{2}$$

where the number of reaction steps (n) in the sequence also equals the number of non-reactant (interior) vertices in the hull subsection.

If every secondary reaction step is required to be the one with the minimum energy (i.e.,

11

largest driving force), then only one unique reaction sequence exists in the left and right hull subsections. However, one must consider the situation where the minimum-energy principle does not hold due to kinetic limitations; this applies in particular to hulls where all secondary reactions have similar magnitude driving forces or a particular phase is kinetically limited from forming, perhaps due to an overall small driving force. Therefore, we must consider the alternative secondary reaction sequences shown in Figures 3b,c. These alternative sequences are not necessarily less favorable; while each alternative sequence may feature a first reaction step with a smaller driving force (i.e., small $\Delta G_{d_{j,1}}$), the latter steps may have larger magnitude energies, resulting in comparable total energy ($\Delta G_{2,j}$) for the particular sequence.

To encompass all combinatorial possibilities in our estimation of secondary competition, we choose to compute the mean total energy of all feasible secondary reaction sequences:

$$\overline{\Delta G_2} = \frac{1}{N} \sum_{j=1}^{N} \Delta G_{2,j}. \tag{3}$$

Determining the total number of unique secondary reaction sequences, N, is mathematically equivalent to calculating the total number of full binary trees with n interior nodes, which yields the Catalan number sequence, $u_n = 1, 1, 2, 5, 14, 42, 132, 429, ...(n = 0, 1, 2, ...)$.[32] Using this connection to the Catalan numbers, we developed a non-recursive algorithm for calculating $\overline{\Delta G_2}$ that is significantly faster than the equivalent recursive solution (see Methods).

Finally, we formulate the C_2 metric such that it accounts for all possible reaction sequences in either hull subsection to the left (L) and right (R) of the target phase:

$$C_2 = - \left(\overline{\Delta G_{2,L}} + \overline{\Delta G_{2,R}} \right). \tag{4}$$

The negative factor is included so that a lower C_2 value corresponds to a more favorable selectivity. Because our definition of a secondary reaction assumes that $\Delta G_d \leq 0$, the secondary competition metric obeys the limiting behavior: $C_2 \geq 0$.

We note that both C_1 and C_2 implicitly assume that the target phase is thermodynamically stable ("on the hull") under the conditions for which the equilibrium phase diagram is derived. However, the competition metrics are still calculable for a metastable phase by manually decreasing its energy until it becomes stable.

Application to Experimental Literature

Using the competition metrics C_1 and C_2, we now assess the selectivities of solid-state reactions previously reported in the experimental literature and use these to rank synthesis recipes by their predicted thermodynamic optimality. Reaction energies and competition metrics were calculated for 3,520 unique experiments reported in the text-mined solid-state reaction literature dataset by Kononova et al.[33] Each unique experiment corresponds to a particular balanced reaction, maximum synthesis temperature, and atmospheric environment (e.g., air, flowing O_2, etc.). We modeled all reactions containing up to two solid precursors and an optional gaseous reactant (see Methods). Reactions that were reported with no particular atmospheric environment are denoted as "closed" and modeled with Gibbs free energies (ΔG_{rxn}), while those with a defined environment are denoted as "open" and modeled using grand potential energies ($\Delta \Phi_{\mathrm{rxn}}$). The results of these calculations are shown in Figures 4a,b.

The synthesis "maps" (Figures 4a,b) allow one to identify favorable reactions by comparing the relative weights of the three reaction metrics: reaction energy (ΔE_{rxn}), primary competition (C_1), and secondary competition (C_2). Thermodynamically optimal reactions are ones that minimize all three metrics, resulting in placement in the lower left region of each plot. According to our calculations, many reactions reported in the literature are predicted to be energetically favorable and selective. Approximately 17.3% of reactions have no exergonic competing reactions on the interface reaction hull (i.e., they are on the bounding line $C_1 = \Delta E_{\mathrm{rxn}}$), and 23.0% of reactions have negligible secondary competition ($C_2 \leq 0.001$ eV/atom). Assuming the literature reactions are experimentally feasible, one would expect
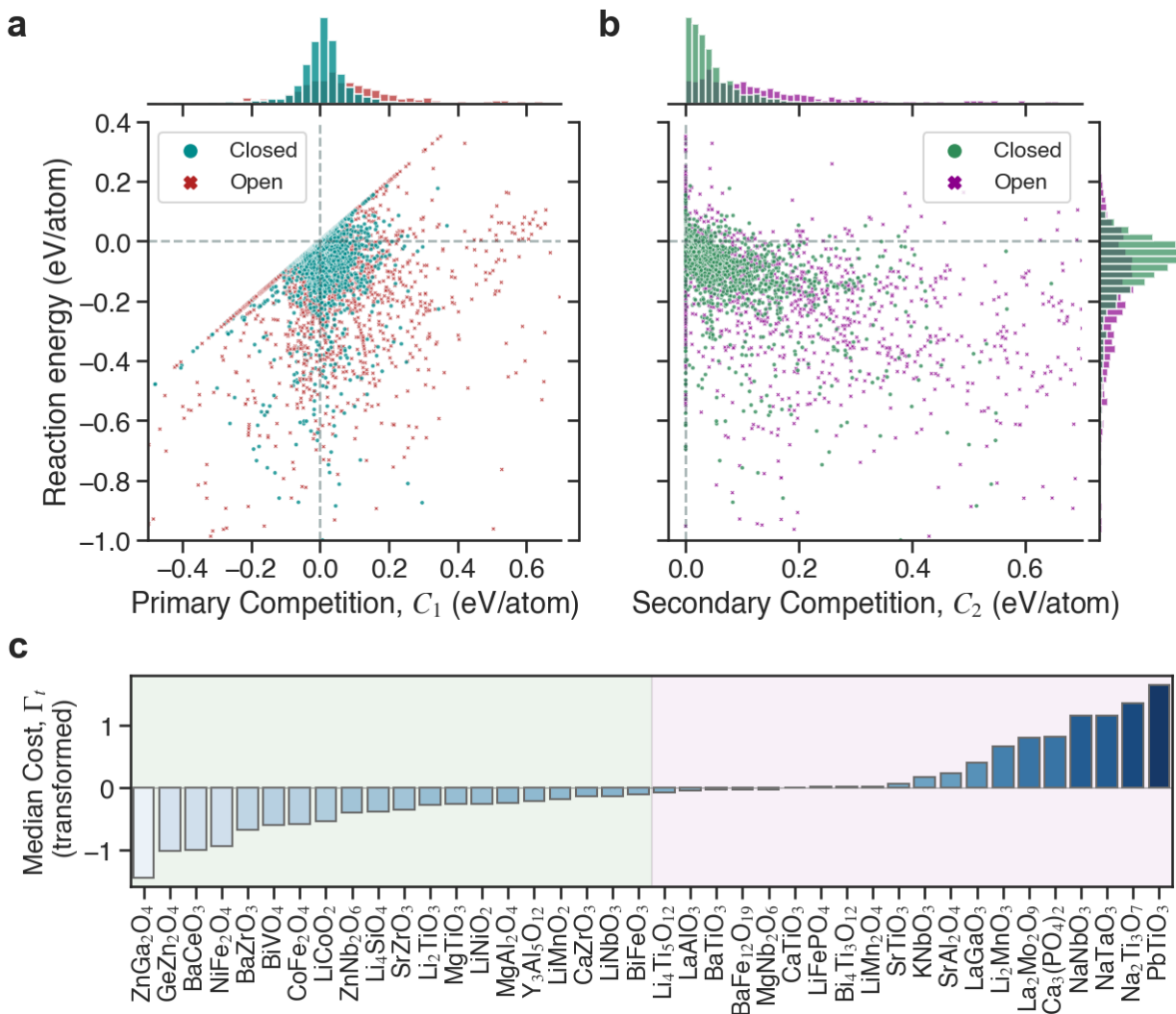
Figure 4: Thermodynamic analysis of synthesis recipes in the experimental solid-state literature. Synthesis maps of 3,520 literature reactions from the Kononova et al. dataset, plotted on a shared axis of reaction energy, ΔE_{rxn}, and independent axes of (a) primary competition, C_1, and (b) secondary competition, C_2. These selectivity metrics are constrained by their lower bounds: $C_1 = \Delta E_{rxn}$ (diagonal parity line) and $C_2 = 0$. (c) Median transformed cost (Γ_t) rankings of synthesis recipes for each of the 40 most popular targets in the dataset. The shading marks the targets with average recipes below (green) or above (pink) the median Γ_t of all experiments in the dataset (-0.089). Selected recipes are discussed in Tables 1 and 2. The full data set is provided in the Supporting Information.

all calculated reaction energies to be negative. Our thermodynamic modeling captures this within reasonable error: 82.8% of reactions have a negative reaction energy, and 97.0% of reactions have $\Delta E_{rxn} \leq 0.1$ eV/atom. Of the reactions with positive energies, most (87.1%)

contain one or more common gases: O_2 (85.1%), CO_2 (48.8%), or H_2O (8.4%). A major source of error in our reaction energy calculations is likely the disagreement between the assumed and actual gas partial pressures; it is often challenging to model the actual environmental conditions of synthesis, as it requires that they be both 1) accurately reported and 2) correctly extracted from the text. Another known source of error is systematic challenges in estimating GGA-calculated formation energies of carbonate compounds.[14] However, this has been addressed and partially mitigated with a fitted energy correction (see Methods).

To quantitatively assess thermodynamic optimality, we follow a similar approach to prior works[24,26] and define a cost function, Γ, that combines the driving force and reaction selectivity evaluated at a particular set of conditions (i.e., temperature and atmosphere). In this work, we opt to use a simple linear weighted summation of the reaction's energy and competition scores,

$$\Gamma = x_0 \Delta E_{\text{rxn}} + x_1 C_1 + x_2 C_2, \tag{5}$$

where ΔE_{rxn} is the reaction energy (either ΔG_{rxn} or $\Delta \Phi_{\text{rxn}}$), and x_0, x_1, and x_2 are user-defined weights for each parameter. Due to the different scaling of each parameter, we find that a non-equal weighting of $x_0 = 0.10$, $x_1 = 0.45$, and $x_2 = 0.45$ produces reasonably diverse results that do not significantly favor one parameter over another. We note that this selection is arbitrary and subject to further optimization.

Unfortunately, closed and open reactions cannot be rigorously compared due to their different energy scales (Gibbs free energies vs. grand potential energies). This is further evidenced by differences in the reaction metric distributions (Figures 4a,b). Since no ideal solution exists for ranking and comparing reactions under different environmental conditions, we transform the cost function to account for the energy scale difference. To do this, we apply a power transformation to the cost distributions for the closed and open reactions, resulting in monotonically transformed costs, Γ_t, whose distributions are closer to standard normal distributions (Figure S1). This new variable facilitates a fairer comparison between closed and open reactions, allowing for a more realistic ranking of synthesis recipes.

Figure 4c shows the median values of Γ_t calculated from the synthesis recipes of the 40 most popular targets in the literature dataset (i.e., those with the most reactions extracted from the text). Given the extent of coverage of our thermodynamic data, we limit our analysis to only the targets for which we have at least five recipes successfully calculated (see Methods). In the following sections, we select several targets with costs below and above the median Γ_t value, analyzing the factors leading to their optimal and suboptimal synthesis recipes, respectively.

Optimal Literature Recipes

Of the 40 most popular targets in the literature dataset (Figure 4c), twenty have synthesis recipes with an average cost value (Γ_t) below the dataset's median (-0.089), indicating generally favorable thermodynamic optimality. A selection of these targets and their highest/lowest-ranked recipes are provided in Table 1, along with other selected reactions of interest. DOIs and raw (untransformed) costs for each reaction are provided in the Supporting Information.

For many targets in Table 1, the conventional reaction involving off-the-shelf binary precursors is predicted to be thermodynamically optimal. For example, in the synthesis of the spinel $ZnGa_2O_4$, the reaction between the binary oxides, $ZnO + Ga_2O_3 \longrightarrow ZnGa_2O_4$, is already perfectly selective ($C_1 = \Delta E_{rxn}$ and $C_2 = 0$) over all temperatures in the dataset. Furthermore, these precursors appear in *all* 17 calculated literature recipes for this target. The favorability of the conventional route seems to apply to several other targets presented here, including $BiVO_4$, $CoFe_2O_4$, $Y_3Al_5O_{12}$, and $BiFeO_3$. Since the precursors are already favorable for these phases, we can instead analyze which conditions are most favorable for target phase formation. For $CoFe_2O_4$, higher temperatures in an open oxygen environment appear to be the most optimal. On the other hand, lower temperatures in an open oxygen environment appear to favor the production of $BiVO_4$ and $BiFeO_3$. Finally, for $Y_3Al_5O_{12}$, there appears to be a tradeoff between selectivity (which is optimal at lower temperatures)

16

and driving force (which is optimal at high temperatures); it appears that intermediate temperatures (600 ℃) in a closed environment result in the most suitable compromise. We note, however, that there are many reasons to use specific processing conditions outside of the pursuit of target phase purity (e.g., improved density, annihilation of defects, optimization of crystallite sizes, etc.). These alternative reasons may explain some of the variability of conditions reported in the literature.

Interestingly, several targets in Table 1 feature synthesis recipes that appear extremely unfavorable. These often involve elemental precursors, such as $Bi + 0.5\ V_2O_5 + 0.75\ O_2 \longrightarrow BiVO_4$, or $Bi + 0.5\ Fe_2O_3 + 0.75\ O_2 \longrightarrow BiFeO_3$. Referencing the original articles from which these reactions were sourced[34,35] suggests that these precursors were not actually used, and the reaction's inclusion in the dataset is likely the result of an error in the literature extraction process. This explains some other impractical and highly suboptimal routes in the dataset, such as reactions involving alkali metal precursors (e.g., $SiO_2 + 4\ Li + O_2 \longrightarrow Li_4SiO_4$).

For other targets such as $LiCoO_2$, Li_4SiO_4, and $LiNiO_2$, there appears to be great variability in precursor selection in the literature. For $LiCoO_2$, the optimal synthesis recipe from our calculations is the reaction of $LiOH$ and Co_3O_4 in a flowing oxygen environment at low to intermediate temperatures (i.e., 700 ℃). The use of Co_3O_4 appears to significantly outperform recipes using $CoCO_3$. Additionally, using $LiOH$ may offer a thermodynamic advantage over Li_2CO_3, although with little effect on the reaction selectivity. A similar conclusion applies to the synthesis of Li_4SiO_4, although the use of lithium carbonate appears to be more favorable at high temperatures in an open-oxygen environment. $LiOH$ (particularly the monohydrate) also appears to offer some advantage in the synthesis of $LiNiO_2$, especially in oxygen at low temperatures (480-600 ℃).

Suboptimal Literature Recipes

Reactions can still be successful despite high C_1 and C_2 values. However, our thermodynamic assessment suggests that these reactions are suboptimal and likely require some combination

Table 1: Thermodynamic analysis of optimal experimental synthesis recipes for selected popular targets in the literature. The ranking of reactions is determined by the transformed cost, Γ_t. The highest- and lowest-ranked reactions are shown along with selected reactions of interest. Raw costs and DOIs are available in the Supporting Information.

Target	Rank	Reaction	Temp. (ºC)	Energy (eV/at)	Open	C_1 (eV/at)	C_2 (eV/at)	Γ_t (a.u.)
$ZnGa_2O_4$	1	$ZnO + Ga_2O_3 \longrightarrow ZnGa_2O_4$	1400	-0.218	O	-0.218	0.000	-1.622
	12	$ZnO + Ga_2O_3 \longrightarrow ZnGa_2O_4$	1200	-0.088		-0.088	0.000	-0.969
	17	$ZnO + Ga_2O_3 \longrightarrow ZnGa_2O_4$	500	-0.067		-0.067	0.000	-0.818
$BiVO_4$	1	$0.5\ V_2O_5 + 0.5\ Bi_2O_3 \longrightarrow BiVO_4$	600	-0.529	O	-0.210	0.171	-1.156
	5	$0.5\ V_2O_5 + 0.5\ Bi_2O_3 \longrightarrow BiVO_4$	500	-0.169		-0.061	0.034	-0.713
	12	$Bi + 0.5\ V_2O_5 + 0.75\ O_2 \longrightarrow BiVO_4$	600	-1.860	O	0.801	4.489	2.412
$CoFe_2O_4$	1	$CoO + Fe_2O_3 \longrightarrow CoFe_2O_4$	1250	-0.088	O	-0.062	0.000	-0.877
	3	$0.3333\ Co_3O_4 + Fe_2O_3 \longrightarrow CoFe_2O_4 + 0.1667\ O_2$	950	-0.058	O	-0.018	0.011	-0.662
	11	$2\ Fe + Co + 2\ O_2 \longrightarrow CoFe_2O_4$	1400	-1.738	O	0.223	2.957	2.111
	14	$CoCO_3 + Fe_2O_3 \longrightarrow CoFe_2O_4 + CO_2$	1400	-0.215		0.156	0.370	2.781
$LiCoO_2$	1	$LiOH + 0.3333\ Co_3O_4 + 0.08333\ O_2 \longrightarrow LiCoO_2 + 0.5\ H_2O$	700	-0.120	O	-0.012	-0.000	-0.727
	4	$0.5\ Li_2CO_3 + 0.3333\ Co_3O_4 + 0.08333\ O_2 \longrightarrow LiCoO_2 + 0.5\ CO_2$	950	-0.061	O	-0.014	-0.000	-0.692
	23	$0.5\ Li_2CO_3 + CoCO_3 + 0.25\ O_2 \longrightarrow LiCoO_2 + 1.5\ CO_2$	300	-0.121	O	0.154	0.357	0.532
	30	$Co + LiOH{\cdot}H_2O + 0.75\ O_2 \longrightarrow LiCoO_2 + 1.5\ H_2O$	900	-0.382	O	1.001	1.454	1.992
Li_4SiO_4	1	$SiO_2 + 2\ Li_2CO_3 \longrightarrow Li_4SiO_4 + 2\ CO_2$	1445	-0.040	O	-0.017	-0.000	-0.683
	7	$SiO_2 + 4\ LiOH \longrightarrow Li_4SiO_4 + 2\ H_2O$	700	-0.049		-0.004	-0.000	-0.412
	10	$SiO_2 + 2\ Li_2CO_3 \longrightarrow Li_4SiO_4 + 2\ CO_2$	1150	-0.010		-0.005	-0.000	-0.366
	18	$SiO_2 + 4\ Li + O_2 \longrightarrow Li_4SiO_4$	800	-2.133	O	0.163	2.830	2.038
$LiNiO_2$	1	$NiO + LiOH{\cdot}H_2O + 0.25\ O_2 \longrightarrow LiNiO_2 + 1.5\ H_2O$	480	0.008	O	0.010	0.002	-0.559
	4	$NiO + 0.5\ Li_2O + 0.25\ O_2 \longrightarrow LiNiO_2$	800	0.032	O	0.058	0.021	-0.344
	6	$Ni(OH)_2 + LiOH + 0.25\ O_2 \longrightarrow LiNiO_2 + 1.5\ H_2O$	770	0.074	O	0.074	-0.000	-0.331
	24	$NiO + 0.5\ Li_2O_2 \longrightarrow LiNiO_2$	800	-0.035		0.067	0.147	0.973
	27	$Ni(OH)_2 + Li + 0.5\ O_2 \longrightarrow LiNiO_2 + H_2O$	700	-0.536	O	1.770	3.044	2.396
$Y_3Al_5O_{12}$	1	$1.5\ Y_2O_3 + 2.5\ Al_2O_3 \longrightarrow Y_3Al_5O_{12}$	600	-0.071		-0.034	0.000	-0.623
	15	$1.5\ Y_2O_3 + 2.5\ Al_2O_3 \longrightarrow Y_3Al_5O_{12}$	1727	-0.240	O	-0.051	0.221	-0.260
	36	$5\ Al(OH)_3 + 1.5\ Y_2O_3 \longrightarrow Y_3Al_5O_{12} + 7.5\ H_2O$	1300	-0.120		-0.007	0.118	0.202
$BiFeO_3$	1	$0.5\ Bi_2O_3 + 0.5\ Fe_2O_3 \longrightarrow BiFeO_3$	600	-0.073	O	0.034	0.053	-0.386
	4	$0.5\ Bi_2O_3 + 0.5\ Fe_2O_3 \longrightarrow BiFeO_3$	100	-0.011		0.000	0.004	-0.312
	38	$Bi + 0.5\ Fe_2O_3 + 0.75\ O_2 \longrightarrow BiFeO_3$	900	-1.285	O	1.087	3.558	2.364

of 1) long heating times to promote thermodynamic equilibrium, 2) follow-up regrinding steps, or 3) fine-tuning of temperature and reaction atmosphere. In Table 2, we highlight several popular target materials that are associated with higher-than-average costs for their synthesis recipes.

Our findings suggest that the targets in Table 2 require a more judicious synthesis due to inherently greater competition in their phase space. Many targets are ranked poorly because the conventional reaction is predicted to be suboptimal. This appears to be true

Table 2: Thermodynamic analysis of suboptimal experimental synthesis recipes for selected popular targets in the literature. The ranking of reactions is determined by the transformed cost, Γ_t. The highest- and lowest-ranked reactions are shown along with selected reactions of interest. Raw costs and DOIs are available in the Supporting Information.

Target	Rank	Reaction	Temp. (°C)	Energy (eV/at)	Open	C_1 (eV/at)	C_2 (eV/at)	Γ_t (a.u.)
PbTiO$_3$	1	$PbCO_3 + TiO_2 \longrightarrow PbTiO_3 + CO_2$	1200	-0.402	O	0.118	0.750	0.960
	2	$PbO + TiO_2 \longrightarrow PbTiO_3$	1100	-0.360	O	0.186	0.694	0.987
	13	$PbO + TiO_2 \longrightarrow PbTiO_3$	800	-0.163		0.136	0.288	2.172
Na$_2$Ti$_3$O$_7$	1	$2\ NaOH + 3\ TiO_2 \longrightarrow Na_2Ti_3O_7 + H_2O$	750	-0.058	O	0.079	0.046	-0.273
	3	$Na_2CO_3 + 3\ TiO_2 \longrightarrow Na_2Ti_3O_7 + CO_2$	1100	0.009		0.108	0.104	1.025
	8	$Na_2CO_3 + 3\ TiO_2 \longrightarrow Na_2Ti_3O_7 + CO_2$	80	0.053		0.202	0.200	2.344
NaTaO$_3$	1	$0.5\ Na_2CO_3 + 0.5\ Ta_2O_5 \longrightarrow NaTaO_3 + 0.5\ CO_2$	1150	-0.088		0.035	0.085	0.300
	4	$0.5\ Na_2CO_3 + 0.5\ Ta_2O_5 \longrightarrow NaTaO_3 + 0.5\ CO_2$	1200	-0.233	O	0.285	0.665	1.098
	9	$0.5\ Na_2CO_3 + 0.5\ Ta_2O_5 \longrightarrow NaTaO_3 + 0.5\ CO_2$	900	-0.194	O	0.441	0.866	1.435
Ca$_3$(PO$_4$)$_2$	1	$CaCO_3 + Ca_2P_2O_7 \longrightarrow Ca_3(PO_4)_2 + CO_2$	800	-0.068		-0.013	-0.000	-0.494
	3	$3\ CaCO_3 + 2\ NH_4H_2PO_4 \longrightarrow Ca_3(PO_4)_2 + 3\ CO_2 + 3\ H_2O + 2\ H_3N$	1150	-0.199	H	0.078	0.269	0.187
	4	$3\ CaCO_3 + 2\ (NH_4)_2HPO_4 \longrightarrow Ca_3(PO_4)_2 + 3\ CO_2 + 3\ H_2O + 4\ H_3N$	1200	-0.156	N	0.105	0.374	0.463
	10	$3\ CaO + P_2O_5 \longrightarrow Ca_3(PO_4)_2$	550	-0.641		-0.007	0.463	1.673
	12	$3\ CaCO_3 + 2\ (NH_4)_2HPO_4 \longrightarrow Ca_3(PO_4)_2 + 3\ CO_2 + 3\ H_2O + 4\ H_3N$	1200	-0.143		0.094	0.336	2.232
Li$_2$MnO$_3$	1	$Li_2CO_3 + MnO_2 \longrightarrow Li_2MnO_3 + CO_2$	975	-0.096	O	-0.015	0.015	-0.670
	4	$2\ LiOH + MnO_2 \longrightarrow Li_2MnO_3 + H_2O$	650	-0.092		-0.024	0.010	-0.533
	5	$MnO_2 + 2\ LiOH \cdot H_2O \longrightarrow Li_2MnO_3 + 3\ H_2O$	70	-0.059		0.001	0.016	-0.295
	14	$MnCO_3 + Li_2CO_3 + 0.5\ O_2 \longrightarrow Li_2MnO_3 + 2\ CO_2$	500	-0.196	O	0.262	0.545	0.942
	15	$MnCO_3 + 2\ LiOH \cdot H_2O + 0.5\ O_2 \longrightarrow Li_2MnO_3 + CO_2 + 3\ H_2O$	450	-0.187	O	0.253	0.554	0.943
	25	$2\ Li + MnO_2 + 0.5\ O_2 \longrightarrow Li_2MnO_3$	700	-1.971	O	0.400	2.399	1.989
LiMn$_2$O$_4$	1	$0.5\ Li_2CO_3 + 2\ MnO_2 \longrightarrow LiMn_2O_4 + 0.5\ CO_2 + 0.25\ O_2$	700	-0.032	O	0.041	0.051	-0.347
	2	$LiOH + 2\ MnO_2 \longrightarrow LiMn_2O_4 + 0.5\ H_2O + 0.25\ O_2$	1000	-0.052	O	0.031	0.067	-0.343
	12	$0.5\ Li_2CO_3 + Mn_2O_3 + 0.25\ O_2 \longrightarrow LiMn_2O_4 + 0.5\ CO_2$	950	-0.073	O	0.030	0.110	-0.240
	47	$2\ MnCO_3 + 0.5\ Li_2CO_3 + 0.75\ O_2 \longrightarrow LiMn_2O_4 + 2.5\ CO_2$	400	-0.319	O	0.148	0.488	0.664
	63	$Mn_2O_3 + 0.5\ Li_2O_2 \longrightarrow LiMn_2O_4$	800	-0.138		0.121	0.283	2.072
LiFePO$_4$	1	$0.3333\ Li_3PO_4 + 0.3333\ Fe_3(PO_4)_2 \longrightarrow LiFePO_4$	600	-0.020		-0.008	0.002	-0.386
	2	$LiPO_3 + 0.5\ Fe_2O_3 \longrightarrow LiFePO_4 + 0.25\ O_2$	1200	-0.151	O	0.108	0.120	-0.060
	3	$0.3333\ Li_3PO_4 + 0.3333\ Fe_3(PO_4)_2 \cdot 8H_2O \longrightarrow LiFePO_4 + 2.667\ H_2O$	800	-0.065		0.000	0.068	0.012
	4	$LiOH \cdot H_2O + Fe(PO_4) \cdot 2H_2O \longrightarrow LiFePO_4 + 3.5\ H_2O + 0.25\ O_2$	650	-0.111	O	0.062	0.247	0.149
	5	$0.5\ Li_2CO_3 + FePO_4 \longrightarrow LiFePO_4 + 0.5\ CO_2 + 0.25\ O_2$	700	0.006	O	0.189	0.122	0.207
BaTiO$_3$	1	$BaCO_3 + TiO_2 \longrightarrow BaTiO_3 + CO_2$	1050	-0.056	O	0.075	0.065	-0.230
	16	$BaCO_3 + TiO_2 \longrightarrow BaTiO_3 + CO_2$	1100	-0.028		0.025	0.026	-0.045
	37	$BaO + TiO_2 \longrightarrow BaTiO_3$	800	-0.239		0.061	0.168	0.778
	40	$BaO + TiO_2 \longrightarrow BaTiO_3$	1300	-0.225		0.080	0.167	0.916

for lead titanate (PbTiO$_3$), which on average, has the most suboptimal recipe of any target investigated (Figure 4c). The high C_1 and C_2 values for PbTiO$_3$ synthesis seem to be

almost entirely related to the instability of the PbO precursor, both with respect to decomposition and to the formation of a solid-solution phase. In our calculations, the latter manifests as the predicted stabilization of a theoretical $Pb_{15}TiO_{17}$ phase. Our modeling of PbO instability supports experimental observations suggesting that the volatility of PbO at higher temperatures results in a challenging $PbTiO_3$ synthesis.[36] The use of $PbCO_3$ as an alternative appears to perform similarly, albeit with a slightly more favorable driving force and lower C_1.

For $Ca_3(PO_4)_2$, Li_2MnO_3, and $LiFePO_4$, the highest-ranked synthesis recipes appear to be nearly thermodynamically optimal already (i.e., low C_1 and C_2). The optimal recipe for $Ca_3(PO_4)_2$ is the reaction of stoichiometric amounts of $CaCO_3$ and $Ca_2P_2O_7$ at moderate temperatures (800 ℃) in a closed environment. For Li_2MnO_3, the reaction of Li_2CO_3 and MnO_2 in open-oxygen environments at moderately high temperatures (900-975 ℃) is favorable, and the use of LiOH in a closed environment at lower temperatures (650 ℃) appears to be similarly advantageous.[37] Finally, for the synthesis of $LiFePO_4$, the use of Li_3PO_4 and $Fe_3(PO_4)_2$ in a closed environment at lower temperatures (600 ℃)[38] appears to be highly selective and greatly outperforms the other recipes analyzed in the literature dataset.

For the remaining targets $Na_2Ti_3O_7$, $NaTaO_3$, $LiMn_2O_4$, and $BaTiO_3$, even the highest-ranked literature synthesis recipes are theoretically suboptimal, which suggests that these materials are suitable candidates for future synthesis optimization efforts. Recipes for $Na_2Ti_3O_7$ generally feature poor driving forces and low selectivities; however, using NaOH as an alternative precursor mitigates this some. All $NaTaO_3$ recipes in our dataset used the same precursors, but those with open-oxygen environments resulted in substantially higher phase competition. $LiMn_2O_4$ synthesis appears to follow similar trends as those discussed previously for $LiMn_2O_3$; the lithium carbonate route in open air appears to be the most optimal of the recipes explored, despite a somewhat low driving force and moderate competition.[39] The use of a $MnCO_3$ precursor is not recommended due to greatly decreased selectivity. Finally, for $BaTiO_3$, the conventional $BaCO_3$ route at high temperatures (1050

°C) and open oxygen conditions appears to be the most favorable; however, with this route, the driving force is still somewhat low and features a moderate amount of phase competition. The $BaTiO_3$ system will be examined extensively in the following sections as an experimental case study for assessing reaction selectivity.

The literature reaction data reveal an inherent tradeoff between driving force and selectivity. Reactions with *only* elemental precursors (e.g., Li metal, O_2 gas) typically have significantly greater driving forces but also much higher competition scores (i.e., lower selectivity), with median values of $\Delta G_{\text{rxn}} = -0.272$, $C_1 = 0.016$, and $C_2 = 0.112$ eV/atom for the closed reactions. For reactions with at least one binary (two-element) precursor, these values shift to $\Delta G_{\text{rxn}} = -0.037$, $C_1 = 0.006$, and $C_2 = 0.020$ eV/atom. In other words, sourcing an element from a precursor with pre-formed bonds yields an interface reaction hull representing a more selective slice of the phase diagram, but at the expense of sacrificing some of the available reaction energy. Fortunately, this tradeoff is not universal and can be circumvented by considering more complex precursor chemistries containing additional elements other than those in the target composition. For example, in oxide synthesis, the use of hydroxides, carbonates, and salts (as in metathesis reactions) often permits the formation of thermodynamically favorable byproducts that increase the reaction's driving force and selectivity. The use of additional elements beyond those found in the target composition has been dubbed "hyperdimensional chemistry" [40] due to its connection to phase diagram geometry; adding a new compositional axis greatly increases the number of ways the phase space can be sliced, allowing one to thermodynamically "shortcut" otherwise unavoidable competing impurity phases. However, not all elemental precursors are unselective, and the necessity of these alternative routes depends on the degree of phase competition in the chemical system of interest. It is typically favorable to synthesize a binary target comprised of elements commonly existing in only a single oxidation state; for example, $Cd + Te \longrightarrow CdTe$ has both perfect selectivity ($C_1 = -0.058$, $C_2 = 0.000$ eV/atom) and a high driving force ($\Delta G_{\text{rxn}} = -0.576$ eV/atom) at 800 °C.

## Case Study: Synthesis of Barium Titanate (BaTiO$_3$)

To investigate whether our reaction selectivity metrics accurately describe impurity formation, we designed and performed experiments testing the influence of precursor selection on the reaction pathway observed during solid-state synthesis. To suggest these experiments, we developed a computational synthesis planning workflow that integrates our proposed selectivity metrics with previous methods for computing solid-state reaction networks.[26] The workflow has five user inputs: 1) target composition, 2) additional elements, 3) temperature, 4) thermodynamic stability cutoff (i.e., maximum E_{hull}), and 5) chemical potential of an open element, if any. The workflow returns a ranked list of possible synthesis reactions based on the cost function implemented in Equation 5.

We selected the ferroelectric barium titanate (BaTiO$_3$) as a case study system for testing our predictions. In addition to being a technologically important and well-studied material in the literature, BaTiO$_3$ is an ideal target for investigating the thermodynamic selectivity of solid-state reactions due to the high number of competing phases in the Ba–Ti–O chemical system (Figure S2). Conventional routes from binary oxides are well-known to proceed through intermediates,[41] and at least 12 unique ternary compositions have been experimentally observed, including the compositional neighbors BaTi$_2$O$_5$ and Ba$_2$TiO$_4$, which frequently appear during synthesis (the latter in particular). Other commonly observed minor impurities include Ba$_4$Ti$_{13}$O$_{30}$, BaTi$_4$O$_9$, and Ba$_6$Ti$_{17}$O$_{40}$. Additional compositions in this space, such as Ba$_2$Ti$_9$O$_{20}$, BaTi$_5$O$_{11}$, Ba$_4$Ti$_{12}$O$_{27}$, and Ba$_2$Ti$_6$O$_{13}$, have been previously synthesized but are less commonly observed as intermediates or impurity phases.[41] Another motivation for selecting BaTiO$_3$ is the kinetic accessibility of its chemical system; even with such a high number of competing Ba–Ti–O phases, the solid-state synthesis of barium titanate is generally considered facile. The ease with which Ba–Ti–O phases can be synthesized suggests that kinetic factors do not significantly divert reaction outcomes from those that are thermodynamically favorable, providing greater justification for our attempts at assessing the likelihood of intermediate/impurity formation through a purely thermodynamic lens.

Using our synthesis planning workflow, we computed and ranked 82,985 binary (two-precursor) synthesis reactions producing $BaTiO_3$, selected from a massive 18-element reaction network composed of all 2,536,160 enumerated binary reactions among 2,417 phases with energies ≤ 0.050 meV/atom above the hull. To capture alternative chemistries (e.g., metathesis reactions, gas-forming reactions, ion exchange reactions, etc.), we selected a chemical system consisting of the target elements (Ba, Ti, O), alkali metals (Li, Na, K) and alkaline earth metals (Mg, Ca, Sr), halogens (F, Cl, Br), chalcogens (S), pnictogens (N, P), and other common elements (B, C, H). The ΔG_{rxn}, C_1, and C_2 values for all calculated synthesis reactions are shown in the synthesis maps illustrated in Figure 5. We calculated these values at an intermediate temperature of $T = 600$ ℃, which is near the median of our experimentally accessible temperature range (see Methods). We also considered the modeling of open-O_2 reactions; similar results are shown in Figure S3 for 62,133 open reactions in a more constrained subsystem.

The determination of an optimal synthesis reaction can be formulated as an optimization problem of simultaneous minimization of the three reaction metrics: ΔG_{rxn}, C_1, and C_2. A common approach for multi-objective optimization in synthesis planning is the identification of the Pareto front.[14] Here, we calculate a three-dimensional Pareto front for the $BaTiO_3$ synthesis reactions (Table 3). The Pareto front reactions for the open-O_2 system are provided in the Supporting Information.

Many Pareto-optimal reactions feature unconventional reactants and byproducts. BaO_2 and $Ba_5(TiN_3)_2$ are the most commonly appearing precursors (eight times each), followed by $Ba(NO_3)_2$ and $Ba_3(PO_4)_2$ (four times each). Nearly all reactions (31 of 33) involve precursors containing additional elements other than Ba, Ti, and O. In particular, nitrogen is used in over half of the reactions (19 of 33), each featuring N_2 gas formation. While unconventional, the high prevalence of nitride precursors in the Pareto front is not theoretically unreasonable; nitrides generally have less negative formation energies than oxides, making oxide formation with N_2 gas evolution both energetically and entropically favorable.
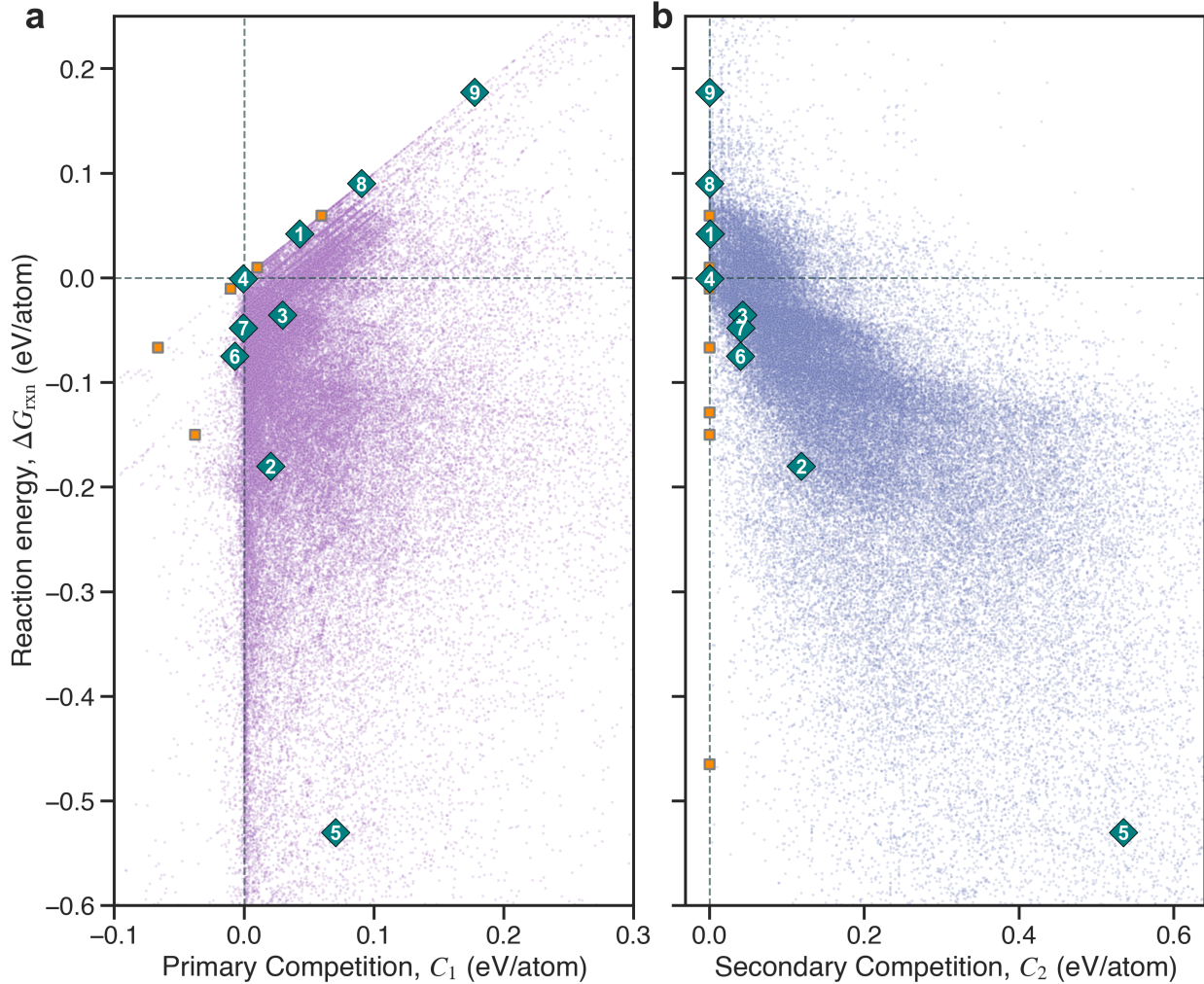
Figure 5: Synthesis maps of 82,985 calculated reactions producing BaTiO$_3$. Target reactions and their competition scores are extracted from an 18-element network of 2,536,160 reactions modeled at a temperature of $T = 600$ ℃. Reactions are plotted on a shared axis of reaction energy, ΔG_{rxn}, and on independent axes of (a) primary competition, C_1, and (b) secondary competition, C_2. The sharp boundaries are lower-bound limits of $C_1 = \Delta G_{\text{rxn}}$ and $C_2 = 0$. Blue diamonds indicate selected reactions experimentally tested in this work. Orange squares represent reactions on the three-dimensional Pareto front of ΔG_{rxn}, C_1, and C_2.

However, many of the reactions appearing on the Pareto front are impractical from an experimental standpoint. For example, the aforementioned nitride precursors are likely challenging to synthesize and handle. Other Pareto front reactions involve theoretical phases (e.g., LiTi$_2$N$_3$), uncommon or toxic precursors (e.g., Ba$_5$(TiN$_3$)$_2$, SO$_2$), difficult-to-remove

Table 3: Pareto front reactions to $BaTiO_3$ and their associated Gibbs free energies, ΔG_{rxn} ($T = 600\ °C$), primary competition scores, C_1, secondary competition scores, C_2, and costs, Γ. All units are in eV/atom.

Rank	Reaction	ΔG_{rxn}	C_1	C_2	Γ	Theoretical
1	$2\ TiP + 4.5\ BaCO_3 \longrightarrow BaTiO_3 + Ba_3(PO_4)_2 + 0.5\ BaTi_2O_5 + 4.5\ C$	-0.465	-0.465	-0.000	-0.256	
2	$3\ Ba(NO_3)_2 + 3.333\ LiTi_2N_3 \longrightarrow BaTiO_3 + 2\ BaTi_2O_5 + 1.667\ Li_2TiO_3 + 8\ N_2$	-0.868	-0.302	-0.000	-0.223	$LiTi_2N_3$
4	$MgTiN_2 + 0.6\ Ba(NO_3)_2 \longrightarrow MgO + 0.2\ BaTiO_3 + 0.4\ BaTi_2O_5 + 1.6\ N_2$	-0.901	-0.223	0.008	-0.187	
6	$8\ O_2 + Ba_4TiP_4 \longrightarrow BaTiO_3 + Ba_3P_4O_{13}$	-2.330	0.062	0.136	-0.144	
7	$4\ BaO_2 + 2\ Ti \longrightarrow Ba_3TiO_5 + BaTiO_3$	-1.489	-0.021	0.036	-0.142	
18	$0.25\ Mg_4TiN_4 + 0.3\ Ba(NO_3)_2 \longrightarrow MgO + 0.2\ BaTiO_3 + 0.05\ Ba_2TiO_4 + 0.8\ N_2$	-0.983	-0.057	0.078	-0.089	Mg_4TiN_4
28	$TiO + BaO_2 \longrightarrow BaTiO_3$	-1.044	-0.034	0.099	-0.075	
33	$0.75\ Ti_7P_4 + 0.3333\ Ba_5P_3O_{12}F \longrightarrow BaTiO_3 + 0.25\ Ba_2TiO_4 + 4\ TiP + 0.1667\ BaF_2$	-0.129	-0.129	-0.000	-0.071	$Ba_5P_3O_{12}F$
54	$MgTi(SO_4)_3 + 4\ BaMg_2 \longrightarrow BaTiO_3 + 9\ MgO + 3\ BaS$	-1.515	-0.005	0.225	-0.052	$MgTi(SO_4)_3$
63	$4\ Ti(NO_3)_4 + 5\ Ba_4P_2O \longrightarrow Ba_2TiO_4 + 3\ BaTiO_3 + 5\ Ba_3(PO_4)_2 + 8\ N_2$	-1.570	-0.018	0.261	-0.047	
85	$2.25\ TiC + 0.5\ Ba_3(PO_4)_2 \longrightarrow BaTiO_3 + 0.25\ Ba_2TiO_4 + TiP + 2.25\ C$	-0.067	-0.067	-0.000	-0.037	
90	$1.25\ Ba_5(TiN_3)_2 + 1.125\ Ti(NO_3)_4 \longrightarrow BaTiO_3 + 2.625\ Ba_2TiO_4 + 6\ N_2$	-1.287	-0.029	0.236	-0.035	
108	$0.75\ Ti_7P_4 + 0.5\ Ba_3(PO_4)_2 \longrightarrow BaTiO_3 + 0.25\ Ba_2TiO_4 + 4\ TiP$	-0.150	-0.038	-0.000	-0.032	
199	$1.45\ Ti(ClO_4)_4 + 0.8\ Ba_6Mg_{23} \longrightarrow BaTiO_3 + 0.45\ Ba_2TiO_4 + 18.4\ MgO + 2.9\ BaCl_2$	-2.540	-0.008	0.527	-0.020	
259	$3\ BaO_2 + Ti_2N_2O \longrightarrow BaTiO_3 + Ba_2TiO_4 + N_2$	-0.913	-0.099	0.268	-0.015	Ti_2N_2O
498	$TiNCl + 1.5\ BaO_2 \longrightarrow BaTiO_3 + 0.5\ BaCl_2 + 0.5\ N_2$	-0.967	-0.150	0.350	-0.007	
547	$Li_2TiO_3 + 0.3333\ Ba_3(PO_4)_2 \longrightarrow BaTiO_3 + 0.6667\ Li_3PO_4$	-0.011	-0.011	-0.000	-0.006	
552	$1.5\ BaO_2 + TiNF \longrightarrow BaTiO_3 + 0.5\ BaF_2 + 0.5\ N_2$	-0.960	-0.153	0.354	-0.006	$TiNF$
651	$TiBrN + 1.5\ BaO_2 \longrightarrow BaTiO_3 + 0.5\ BaBr_2 + 0.5\ N_2$	-0.963	-0.152	0.357	-0.004	
1850	$Ba_3(PO_4)_2 + Ca_4Ti_3O_{10} \longrightarrow Ca_4P_2O_9 + 3\ BaTiO_3$	0.010	0.010	-0.000	0.005	
2825	$0.05882\ Ba_2Mg_{17} + 0.1118\ Ti(NO_3)_4 \longrightarrow MgO + 0.1059\ BaTiO_3 + 0.005882\ Ba_2TiO_4 + 0.2235\ N_2$	-1.896	-0.004	0.447	0.010	
6598	$3\ Ti(SO_4)_2 + 8\ BaMg_2 \longrightarrow BaTiO_3 + 16\ MgO + 6\ BaS + BaTi_2O_5$	-1.628	-0.006	0.413	0.020	
7657	$2\ BaO_2 + 0.25\ NaTi_5(NCl)_5 \longrightarrow BaTiO_3 + 0.25\ Ba_2TiO_4 + 0.25\ NaCl + 0.5\ BaCl_2 + 0.625\ N_2$	-0.932	-0.079	0.336	0.023	$NaTi_5(NCl)_5$
8896	$0.5\ Ba_5(TiN_3)_2 + 3\ KNOF_2 \longrightarrow BaTiO_3 + 3\ KF + 1.5\ BaF_2 + 3\ N_2$	-1.170	-0.040	0.356	0.025	
13191	$TiO_2 + 0.5\ Ba_2Ca(BO_2)_6 \longrightarrow BaTiO_3 + 0.5\ Ca(B_3O_5)_2$	0.060	0.060	-0.000	0.033	
15339	$0.5\ Ba_5(TiN_3)_2 + 3\ NOF \longrightarrow BaTiO_3 + 1.5\ BaF_2 + 3\ N_2$	-1.667	-0.058	0.509	0.036	
39153	$0.5\ Ba_5(TiN_3)_2 + 1.5\ SO_2 \longrightarrow BaTiO_3 + 1.5\ BaS + 1.5\ N_2$	-1.102	-0.042	0.450	0.074	
45245	$0.5\ Ti_2S + 1.5\ BaO_2 \longrightarrow BaTiO_3 + 0.5\ BaS$	-1.366	-0.068	0.563	0.086	
57844	$0.8\ Ba_5(TiN_3)_2 + 0.45\ Ti(ClO_4)_4 \longrightarrow BaTiO_3 + 1.05\ Ba_2TiO_4 + 0.9\ BaCl_2 + 2.4\ N_2$	-1.729	-0.077	0.728	0.120	
64696	$5\ MgTiH_4 + 3\ Ba(NO_3)_2 \longrightarrow BaTiO_3 + 5\ MgO + 10\ H_2 + 2\ BaTi_2O_5 + 3\ N_2$	-1.142	-0.078	0.659	0.147	
65291	$0.8\ Ba_5(TiN_3)_2 + 1.8\ Br_2O_3 \longrightarrow BaTiO_3 + 0.6\ Ba_2TiO_4 + 1.8\ BaBr_2 + 2.4\ N_2$	-1.810	-0.055	0.790	0.150	
65313	$0.5\ Ba_5(TiN_3)_2 + 1.5\ BrO_2F \longrightarrow BaTiO_3 + 1.5\ BaBrF + 1.5\ N_2$	-1.878	-0.053	0.804	0.150	
66253	$3.5\ Ba_5(TiN_3)_2 + 9\ ClO_3 \longrightarrow BaTiO_3 + 6\ Ba_2TiO_4 + 4.5\ BaCl_2 + 10.5\ N_2$	-1.841	-0.021	0.775	0.155	

byproducts (e.g., Ba_2TiO_4), or refractory precursors (e.g., TiC). Some of these suggested reactions can be removed easily with user-applied filters that account for specific experimental

restrictions. For example, one can supply a list of available precursor compositions ("off-the-shelf" phases) or composition types to be avoided (e.g., sulfides, acids, etc.). Other reactions require a more careful manual inspection to validate experimental practicality, particularly with regard to refractory precursors. This challenge can be mitigated by the future inclusion of data or models relevant to solid-state reaction kinetics (e.g., defect formation energies) and the prediction of physical properties (e.g., melting points, hardness).[42] The unorthodox nature of the unprocessed reactions is an advantage of our approach, however, and permits synthesis recommendations that expand beyond traditional chemical intuition.

We selected nine reactions from Figure 5 to test experimentally. The calculated thermodynamic metrics for these reactions are provided in Table 4. We intentionally selected reactions spanning various precursor chemistries, free energies, and competition scores. To prioritize the study of impurity-forming reactions and avoid the aforementioned practicality challenges, we did not explicitly include any reactions on the Pareto front. The conventional synthesis route, $BaCO_3 + TiO_2 \longrightarrow BaTiO_3 + CO_2$, was chosen as a baseline reference (Expt. 1). A positive energy is calculated for this reaction ($\Delta G_{rxn} = +0.042$ eV/atom at $T = 600\,°C$); however, this is likely due to residual uncorrected error in the calculated energy of $BaCO_3$, which is not included in the NIST-JANAF dataset. We did not test the analogous reaction with BaO precursor due to its hygroscopic nature, which makes it difficult to handle. However, we did test the alternative reaction from barium peroxide, BaO_2 (Expt. 2). We included two reactions that form $BaTiO_3$ directly from at least one other ternary phase (Expts. 3, 4). A reaction with a Ti metal precursor was selected due to its extremely high C_2 score (Expt. 5). Two metathesis reactions (Expts. 6, 7) were selected due to their predicted high performance, including the unconventional use of a sulfide precursor (BaS). The final two reactions (Expts. 8, 9) were selected for being endergonic ($\Delta G_{rxn} > 0$) to validate the accuracy of our free energy predictions. We note that several experiments feature precursors that are not easily purchasable from a chemical supplier (e.g., Ba_2TiO_4, Na_2TiO_3); these phases were synthesized following recipes reported in the literature (see Methods).

Table 4: Selected experimental $BaTiO_3$ synthesis reactions and their associated Gibbs free energies, ΔG_{rxn} ($T = 600\,°C$), primary competition scores, C_1, secondary competition scores, C_2, and costs, Γ.

Expt.	Reaction	ΔG_{rxn} (eV/at)	C_1 (eV/at)	C_2 (eV/at)	Γ (eV/at)
1	$BaCO_3 + TiO_2 \longrightarrow BaTiO_3 + CO_2$	0.042	0.043	0.000	0.024
2	$BaO_2 + TiO_2 \longrightarrow BaTiO_3 + 0.5\ O_2$	-0.180	0.021	0.119	0.045
3	$Ba_2TiO_4 + TiO_2 \longrightarrow 2\ BaTiO_3$	-0.036	0.030	0.043	0.029
4	$Ba_2TiO_4 + BaTi_2O_5 \longrightarrow 3\ BaTiO_3$	-0.001	-0.001	0.000	-0.001
5	$Ba(OH)_2 \cdot H_2O + 3.666\ Ti \longrightarrow BaTiO_3 + 1.333\ Ti_2H_3$	-0.530	0.070	0.534	0.219
6	$BaCl_2 + Na_2TiO_3 \longrightarrow BaTiO_3 + 2\ NaCl$	-0.075	-0.007	0.040	0.007
7	$BaS + Na_2TiO_3 \longrightarrow BaTiO_3 + Na_2S$	-0.048	-0.001	0.041	0.013
8	$2\ BaS + 3\ TiO_2 \longrightarrow 2\ BaTiO_3 + TiS_2$	0.090	0.090	0.000	0.050
9	$BaSO_4 + 2\ TiO_2 \longrightarrow BaTiO_3 + TiOSO_4$	0.178	0.178	0.000	0.098

The nine synthesis experiments were completed using a gradient furnace (see Methods),[43] allowing the observation of reaction products over a wide range of temperatures ($\sim$200-1000 ℃) with *ex post facto* SPXRD. The experimental results are summarized in Figure 6. Mole fractions of each phase were determined via Rietveld refinement of SPXRD patterns acquired at various positions (temperatures) along the length of the sample after heating and subsequent cooling to ambient temperature. The *ex post facto* phase fraction plots can be interpreted similarly to those constructed with *in situ* data, as each effectively illustrates the reaction pathway during heating. More rigorously, however, each point in Figure 6 is pseudo-independent and best interpreted as the result of an isothermal (*ex situ*) reaction at its associated temperature. Shorter reaction times were selected to ensure capture of the onset of short-lived intermediate phases critical to assessing reaction pathway selectivities (Table S2). Selected Rietveld analysis results for each experiment are provided in the Supporting Information (Figures S4-S12).

The observed reaction pathways demonstrate significant variation in target and impurity formation. Visualizing the interface reaction hulls of the selected experiments helps to rationalize their predicted and observed performance (Figure S13). The most complex of these pathways is that of Expt. 5, which features the formation of $BaTiO_3$ at an intermediate temperature range (400-500 ℃) before impurities Ba_2TiO_4, TiH_2, and TiO_2 begin to
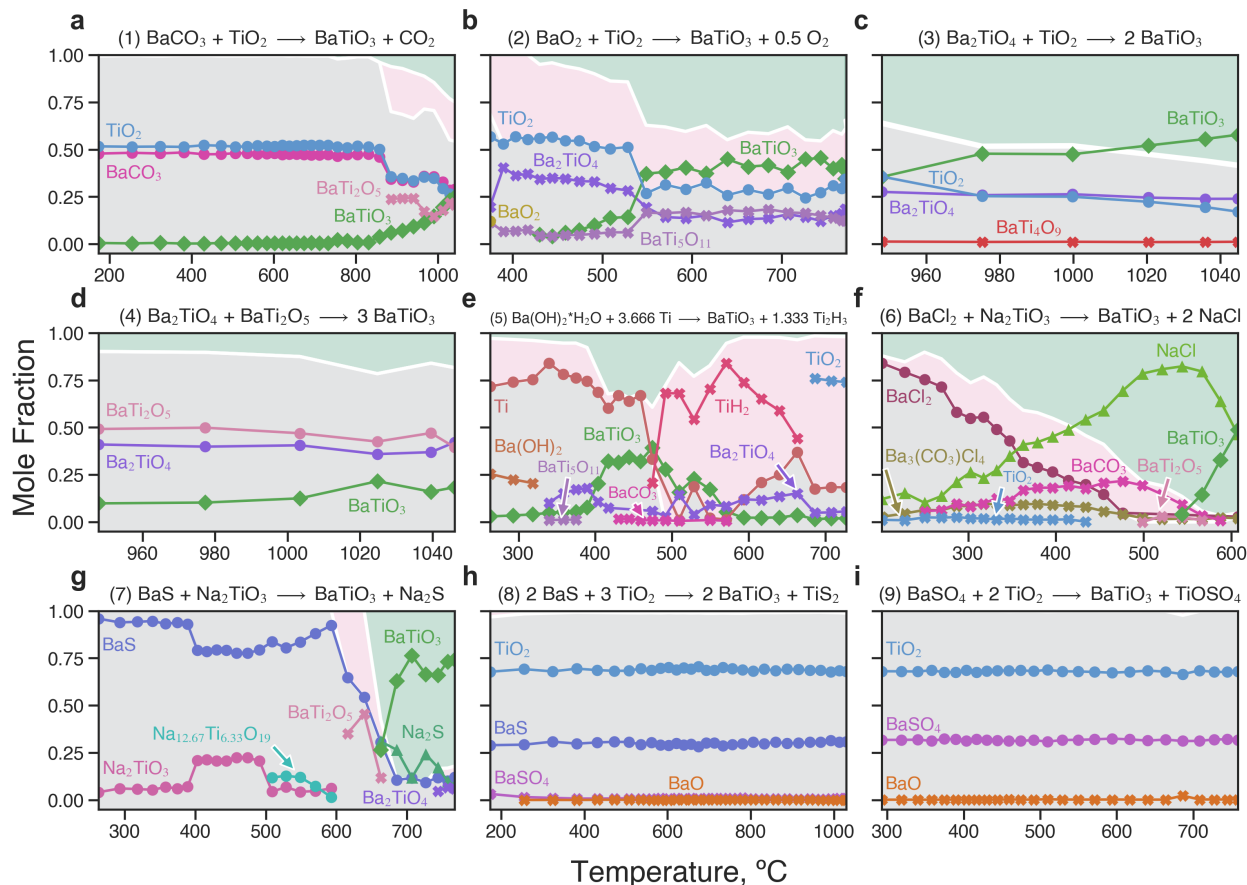
Figure 6: Reaction pathways for selected BaTiO$_3$ synthesis experiments. (a-i) Mole fractions of observed phases for Experiments 1-9, as determined through Rietveld refinements of *ex post facto* SPXRD data. Phase types are distinguished by shape: precursors (circles), targets (diamonds), byproducts (triangles), and impurities (exes). Background shading denotes the total mole fractions of precursor (gray), impurity (pink), and target/byproduct (green). The median total reaction time was ~67 min; exact times for each experiment are provided in Table S2.

dominate. Indeed, Expt. 5 exhibits both the largest driving force (-0.530 eV/atom) and C_2 value (0.534 eV/atom) of any reaction, supporting the observation of a complex reaction pathway containing many impurity phases.

The conventional synthesis reaction between BaCO$_3$ and TiO$_2$ (Expt. 1) was largely incomplete after 60 min at 1000 °C and exhibited significant formation of BaTi$_2$O$_5$. The interface reaction hull for this system (Figure S13a) suggests that the formation of BaTi$_2$O$_5$ is the most favorable reaction outcome, supporting our observations. In Expt. 2, the reaction

of TiO_2 with BaO_2 appears to significantly decrease the reaction onset temperature but also features substantial impurity formation. In three of the experiments (Expts. 4, 6, and 7), the $BaTiO_3$ synthesis reaction is the most favorable reaction on the hull, resulting in $C_1 < 0$. Notably, the use of ternary precursor(s) in Expts. 3 and 4 (Ba_2TiO_4, $BaTi_2O_5$) results in low (or zero) C_2, but also very little driving force. As a result, we observe near-perfect selectivity (i.e., very few visible impurities) at the expense of slowing down the reactions substantially. When the reaction energy is above zero, impurities associated with exergonic competing reactions may still form. For example, the hull for Expt. 8 indicates a significant degree of competition, including several reactions with only slightly positive energies ($\Delta G_{\text{rxn}} < 0.01$ eV/atom). Experiments 8 and 9 are indeed largely unreacted as predicted but feature minor impurities ($BaSO_4$ and/or BaO).

The metathesis reactions (Expts. 6, 7) show the overall greatest performance, yielding primarily $BaTiO_3$ and the predicted byproducts at moderately low temperatures (600-700 ℃). While metathesis reactions producing alkali halides are well-known for their optimal performance,[44] it is a notable and surprising result that the sulfide-based reaction ($BaS + Na_2TiO_3 \longrightarrow BaTiO_3 + Na_2S$) achieves such pure and direct synthesis of $BaTiO_3$, as predicted. Its success further highlights the importance of considering more complex chemistries involving additional elements besides those in the target phase (i.e., hyperdimensional chemistries).[40] Some impurity formation, however, is evident in both metathesis reactions, particularly at lower temperatures. $BaTi_2O_5$ forms in both experiments and small amounts of Ba_2TiO_4 form in Expt. 7. However, this observation is supported by the calculated C_1 and C_2 scores, which indicate that neither reaction should be perfectly selective. Unexpectedly, the dominant impurities in Expt. 6 are carbonate compounds: $Ba_3(CO_3)Cl_4$ and $BaCO_3$. We presume this results from the reaction of the original precursors with CO_2 in the air. For completeness, we have accounted for these in our experimental analysis even though they were not explicitly considered in the selectivity calculations. However, we did exclude from consideration any Si-containing impurities such as $Ba_2TiSi_2O_8$, which formed

in small amounts due to reaction with the quartz capillaries; these Si impurities were minor (<0.02 mole fraction) and did not significantly affect our analysis.

To quantitatively assess the performance of our predictions in determining the outcomes of experimental reactions, we propose three reaction outcome metrics summarizing the behavior of a reaction pathway: the minimum precursor remaining (P), the maximum target/byproduct formed (T), and the maximum impurity formed (I). Each metric is a mole fraction value taken from any data point (i.e., temperature) within the reaction pathway. This permits the capture of key features of the pathway independent of the kinetics of that reaction and is necessary given the range of chemistries explored. Our outcome metrics are visualized on the Precursor-Target-Impurity (PTI) plots in Figure 7a. The first quantity, P (minimum precursor remaining), gives insight into the reactivity and kinetics of the reaction: high values indicate reactions that did not complete at any temperature within the reaction timeframe. The second quantity, T (maximum target/byproduct formed), provides a measure of the success of the reaction in producing $BaTiO_3$ and predicted byproduct(s). Finally, the third quantity, I (maximum impurity formed), measures the selectivity of the reaction pathway, indicating the maximum fraction of intermediate/impurity phases synthesized at any temperature. Note that the temperature-independent PTI metrics are not required to sum to one for a particular experiment. This is intentional and advantageous because it permits the capture of poor selectivity in even nominally well-performing reactions; i.e., both T and I can be high (~ 1) within the same experiment.

Figures 7b-d show selected correlations between reaction outcomes (P, T, and I) and the calculated reaction metrics (ΔG_{rxn}, C_1, and C_2). The full set of all (3x3) pairwise correlation plots is available in Figure S14. We observe that the calculated reaction energy (ΔG_{rxn}) correlates most strongly with the minimum amount of precursor remaining (P) at the conclusion of the experiment (Figure 7b). With infinite reaction times, we would theoretically expect this distribution to resemble a step function: $P = 0$ for reactions with $\Delta G_{\mathrm{rxn}} < 0$, and $P = 1$ for those with $\Delta G_{\mathrm{rxn}} > 0$. In our work, the distribution is less
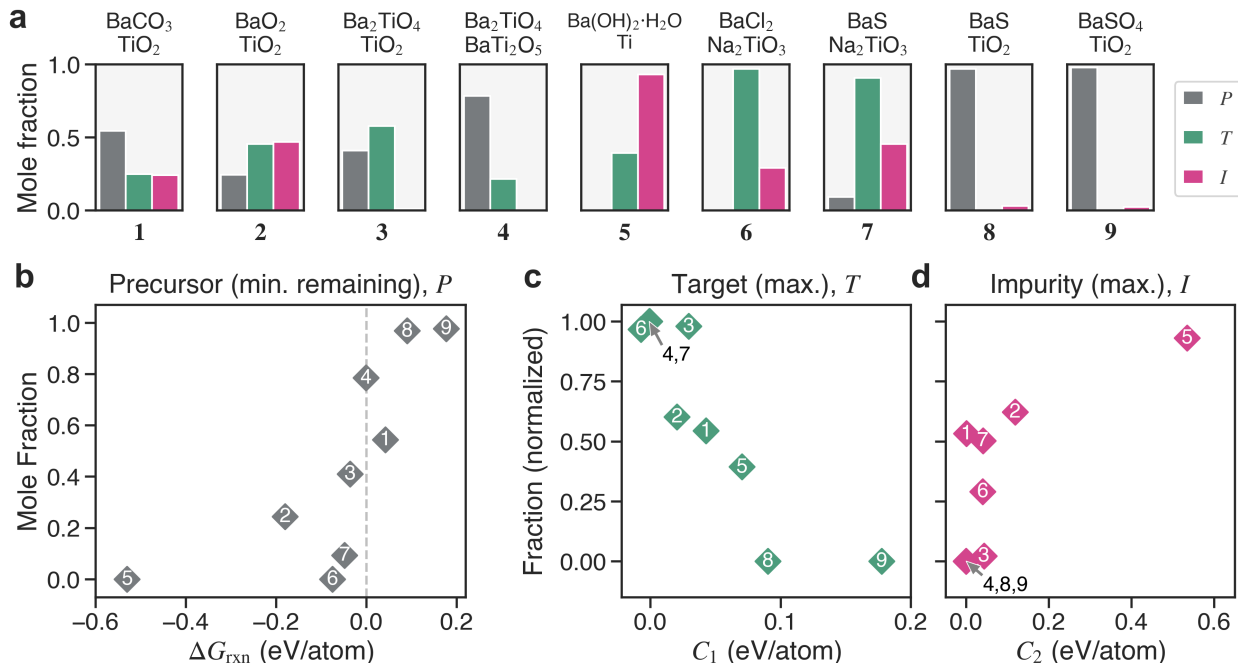
Figure 7: Summary of experimental results and correlations with calculated reaction metrics. (a) Precursor-Target-Impurity (PTI) plots, where the height of each bar indicates the relevant mole fraction captured at the most representative position (i.e., temperature) along the length of the sample after heating (and cooling). The different color bars correspond to the minimum amount of precursor remaining (P, gray), the maximum amount of target/byproducts formed (T, green), and the maximum amount of impurity formed (I, pink). (b) Positive correlation between P and the Gibbs free energy of the reaction, ΔG_{rxn}. (c) Negative correlation between T and C_1. (d) Positive correlation between I and C_2. The T and I mole fractions have been normalized by the maximum amount of precursor consumed in the experiment, $1 - P$, for enhanced visualization of the trend. The small impurity amounts detected in Expts. 8 and 9 are treated as yielding $I = 0$ to more accurately reflect the lack of observed reaction.

defined, given the shorter reaction times and different chemistries explored. The coordinates of Expts. 1 and 2 appear to deviate the most from a step-like distribution. Difficulty in modeling the energetics of carbonate reactions was previously discussed and likely explains the deviation of Expt. 1. In contrast, the deviation of Expt. 2 is likely kinetic in nature, as the reaction appears to stall (Figure 6b); this may suggest that the maximum temperature (750-800 ℃) is too low to achieve sufficient reaction completion.

The primary (C_1) and secondary (C_2) competition metrics display negative and positive

31

correlations with the maximum amounts of target (T) and impurity (I) phases formed, respectively (Figures 7c,d). We note that these trends are not strictly obeyed in a monotonic fashion. Still, the correlations are significant and can be theoretically rationalized by their derivation from the interface reaction hull. It is reasonable that C_1 correlates most strongly with T, as this competition metric effectively measures the relative favorability of the target reaction over competing reactions. Similarly, it is sensible that C_2 should correlate most strongly with I given its derivation as a measure of the *relative* stability of impurity phases with respect to the target phase. To be precise, the experimental correlation between C_2 and I suggests that one should consider not only the sum of the inverse hull distance energies for all competing phases but the *total* energy of the entire secondary reaction sequence containing them (Equation 2). By definition, this quantity includes, and thus will always be larger than or equal to, the sum of all competing inverse hull distance values for a particular interface reaction hull. Regarding the functional form of C_2, the question arises as to whether a more simple summation of the *maximum* energy secondary reactions in the left and right hull subsections is a sufficient measure for secondary competition. While the maximum energy secondary reactions tend to account for much of the value of C_2, on average, the full C_2 metric is 0.077 eV/atom greater than considering the maximum energy secondary reactions alone (Figure S15), suggesting that C_2 is a more conservative metric. For completeness, we also tested an alternative formulation of the secondary competition metric using the enclosed "area" to the hull. While this metric correlates with C_2, its calculation is more numerically unstable, and its units are less interpretable. Hence, we generally recommend the approach of modeling full secondary reaction sequences, which is straightforward to implement using our secondary competition algorithm (see Methods).

In a previous study,[24] we suggested a solid-state reaction selectivity metric based on the difference in elemental chemical potentials between precursors and targets, measured by a distance along the chemical potential diagram (i.e., the "total chemical potential distance"). This metric was used to rationalize the unique selectivity of the Na-based precursor in synthe-

sizing pyrochlore $Y_2Mn_2O_7$ from YOCl and $AMnO_2$ (A=Li, Na, K). While straightforward to compute using just the chemical potential diagram, the distance metric operates in the space of chemical potentials rather than reaction energies, rendering it less intuitive and more difficult to precisely discern the specific competing reactions. From the results here, we generally recommend that C_2 be used instead of total chemical potential distance where possible. Both selectivity metrics capture similar characteristics of the competing phase space: each effectively involves the summation of competing phase stabilities through inverse hull distances or the corresponding chemical potential stability ranges. More precisely, both quantities are correlated (Figure S16) because chemical potentials are mathematical derivatives of the convex hull in energy-composition space. However, the total chemical potential distance is biased, particularly by competing phases with defective elemental-like compositions (e.g., $Mg_{149}Cl$). Due to the increased weight of the entropic ($-TS$) term in the definition of Gibbs free energy, chemical potential diagrams featuring these compositions as competing phases may yield very high (unfavorable) total chemical potential distance values for synthesis reactions.

We acknowledge that while C_1 and C_2 are meant to capture different, independent mechanisms by which competing phases form, these metrics are at least *partially* correlated due to the geometric constraints of the convex hull. In particular, one situation is geometrically limited from occurring: high C_2 and low C_1 (Figure S17). Stated explicitly: if a competing phase lies significantly below the tie-line formed by the target and a precursor (i.e., high C_2), then both the target and that competing phase necessarily have similar reaction energies to form from the precursors, leading to high C_1. In general, however, this restriction does not make the metrics redundant; while there is some correlation between the two, the correlation is not particularly strong (Figure S16). Therefore, we generally recommend the tandem use of both selectivity metrics.

The major limitation of our current synthesis planning workflow is the assumption that optimal synthesis reactions can be predicted with the thermodynamic energy landscape

alone. While this is not the case for all chemistries, we show here that at least for chemical systems that exhibit practical solid-state reaction kinetics, the energy landscape alone can provide much of the rationalization for the observation of impurity phases. To say that impurity and secondary phases are inherently "kinetic" products is a misnomer. Rather, these phases may be the thermodynamic minima of smaller "local" interface systems, distinct from the thermodynamic products of the entire reaction mixture (the global thermodynamic solution). Furthermore, these impurities are often not easily convertible to final products without long-range mass transport or intervention (e.g., via re-grinding or subsequent heating). This explains why impurities are often pervasive and challenging to remove in chemical systems with lower driving forces and/or slower kinetics (e.g., $BiFeO_3$).

Currently, our synthesis planning workflow focuses exclusively on optimizing product purity; however, there are many issues one must consider when designing a synthesis recipe for a target compound: material cost, safety concerns, stability in air, handling challenges, availability of precursors, etc. Many routes suggested by our workflow involve the formation of byproducts that are not easily removable from the product mixture (e.g., the formation of $BaTiO_3$ with byproduct Ba_2TiO_4). To this point, one should be thoughtful in designing criteria by which to filter recommended synthetic routes. For example, one can prioritize the formation of only gaseous byproducts (e.g., O_2 or CO_2) or those easily removable by a solvent (e.g., NaCl). The cost function used to rank reactions can be modified to include other reaction metrics of interest, such as the estimated economic cost of the precursor materials. While not explicitly demonstrated here, the synthesis planning workflow can also be extended to use for multi-step syntheses, allowing one to retrosynthetically sequence reactions to a target material beginning with purchasable, "off-the-shelf" precursors.

Conclusions

Using the interface reaction model for powder reactions, we proposed two thermodynamic selectivity metrics for solid-state reactions: primary (C_1) and secondary (C_2) competition. To systematically and critically examine the effectiveness of our metrics, we analyzed existing successful synthesis routes available in the literature and, leveraging a massive set of 82,985 synthesis reactions extracted from an 18-element reaction network constructed from Materials Project data, designed and executed nine $BaTiO_3$ synthesis experiments with a range of selectivity values as compared to conventional precursors ($BaCO_3$ and TiO_2). Analysis of reaction pathways in the nine experiments via *ex post facto* synchrotron powder X-ray diffraction reveals that C_1 and C_2 correlate with the maximum amounts of target and impurity formed, respectively.

The main advantage of our approach compared to recent, existing approaches[14,25] is the ability to simultaneously consider a wide range of chemistries, including those with unconventional additional elements. These so-called hyperdimensional chemistries[40] allow one to bypass commonly encountered intermediates in target systems with many competing phases. This was demonstrated particularly for the $BaTiO_3$ system studied in this work and is relevant for many other materials in the literature that are conventionally synthesized with theoretically suboptimal precursors (e.g., $Na_2Ti_3O_7$, $NaTaO_3$, $LiMn_2O_4$, etc.).

We anticipate that the selectivity metrics presented here and our computational synthesis planning workflow will significantly reduce the synthesis bottleneck, providing more rapid development of synthesis approaches for new, predicted materials. Our workflow provides a theoretical rationale for using certain precursors and synthesis conditions over other options, which promises to optimize existing synthesis procedures for current technologically important materials.

We envision our approach to be particularly useful in aiding high-throughput automated laboratory exploration efforts.[45] Predictions can be used to design and downselect the synthesis reactions tested, reducing the cost and current trial-and-error approach to inorganic

materials synthesis. The future inclusion of models for the kinetic behavior of reactions, such as estimates of the reactivity of precursors based on solid-state diffusivities, will further enhance predictions.

Methods

Thermodynamic Data

Gibbs free energies of formation, $\Delta G_\mathrm{f}(T)$, were acquired or approximated in a similar approach to those of previous works.[24,26] We acquired experimental $\Delta G_\mathrm{f}(T)$ values from the NIST-JANAF thermochemical tables[29] where available. Experimental values were limited to compounds with low melting points (i.e., $T_m \leq 1500$ ℃), as these systems demonstrate more complex phase change behavior over the temperature range studied here. For predominantly solid compounds (i.e., those with melting points above this threshold), as well as for all other phases not available in the NIST-JANAF thermochemical tables, we estimated $\Delta G_\mathrm{f}(T)$ using the machine-learned Gibbs free energy descriptor identified by Bartel et al.[28] This descriptor was applied using formation enthalpies $\Delta H_\mathrm{f}(T = 298$ K$)$ acquired from the Materials Project (MP) database,[27] version 2022.10.28.

Due to the well-known and systematic formation energy error of carbonate compounds calculated with GGA exchange-correlation functionals,[14,19] we applied an energy correction of 0.830 eV per CO_3^{2-} anion to all carbonate compounds acquired from MP. This value was determined by fitting the mean error between computed and experimental $\Delta G_\mathrm{f}(T = 300$ K$)$ values for 15 metal carbonate compounds (Figure S18).

Synthesis Planning Workflow

The synthesis reaction calculation and ranking procedure was implemented as a Python-based workflow in the existing *reaction-network* package.[26] The code is available on GitHub at `https://github.com/materialsproject/reaction-network`. The workflow was con-

structed and launched on computing resources using the *jobflow*[46] and *fireworks*[47] workflow packages.

The synthesis planning workflow consists of three sequential steps. First, phases and their formation energies for the chemical system of interest are acquired as previously described. The total number of phases can be optionally reduced by setting a threshold for the maximum energy above hull (ΔG_{hull}). In this work, we used a moderately large threshold of $\Delta G_{\text{hull}} \leq 50$ meV/atom, evaluated at ambient temperature ($T = 300$ K). Second, reaction enumeration is performed for the acquired phases using the combinatorial and free energy minimization approaches described in our previous work on solid-state reaction networks.[26] Note that the combinatorial approach allows one to identify reaction product combinations above the hull (i.e., "metastable" products), which makes the analysis more robust to numerical error in the thermodynamic data. For systems with an open element (e.g., O_2 gas), this reaction enumeration step is performed again using grand potential energies, where the open element has been assigned a user-defined value for the chemical potential (often the standard state, $\mu = \mu^0$ eV). Finally, C_1 and C_2 scores are calculated for all target synthesis reactions (i.e., those that form the desired target composition). To do this, the relevant competing reactions are extracted from the full set of enumerated reactions. We define a competing reaction as one whose precursors are a subset of the target reaction's precursors. These competing reactions are then used to compute the interface reaction hull, from which C_1 and C_2 are calculated via Equations 1 and 4. For open systems, this selectivity calculation procedure is performed again, including any additional enumerated open reactions and ensuring that all reactions are calculated with grand potential energies at the corresponding chemical potential.

Secondary Competition Algorithm

The secondary competition score, C_2, is defined as the negative sum of the mean secondary reaction sequence energies to the left and right of the target on the interface reaction hull (Equation 4). One approach for acquiring these quantities involves using a recursive algo-

rithm to identify all possible sequences and their energies. However, this strategy is too slow for the high-throughput calculation of C_2 in systems with many competing reactions.

Instead, we have identified a non-recursive algorithm that takes advantage of the connection between this problem and the recursive construction of binary trees via the use of the Catalan number sequence. Our algorithm reformulates the sum of all secondary reaction sequence energies as a sum of individual secondary reaction energies weighted by their multiplicities, i.e., the total number of appearances of a particular reaction within the set of all possible secondary reaction sequences. The energy of any reaction indexed k can be calculated geometrically as the altitude, h_k, of the triangle formed by its product vertex and two reactant vertices on the interface reaction hull. We find that the altitude multiplicity, m_{h_k}, is determined to be the product of three Catalan numbers, u_n, such that

$$m_{h_k} = u_{n_l} \cdot u_{n_r} \cdot u_{(n-n_l-n_r-1)}, \tag{6}$$

where n_l and n_r refer to the number of interior vertices (i.e., within the triangle) to the left and right of the vertex of interest, respectively, and n is the total number of interior vertices for the entire hull subsection. For example, for secondary reactions between nearest neighbors, $n_l = 0$ and $n_r = 0$, resulting in an altitude multiplicity of $m_h = u_{n-1}$.

The mean secondary reaction sequence energy for the hull subsection can then be calculated as

$$\overline{\Delta G_2} = \frac{1}{N} \sum_{k}^{V} m_{h_k} \cdot h_k, \tag{7}$$

where the sum occurs over all of the V unique reaction energies (altitudes), which is the number of unique triangles that can be constructed for the hull subsection, including the two exterior vertices: $V = \binom{n+2}{3}$. The total number of unique secondary reaction sequences equals the corresponding Catalan number, $N = u_n$. Finally, once this process has been performed for both the left and right hull subsections, the secondary competition (C_2) can be calculated via Equation 4.

Literature Reactions

Solid-state literature reactions studied in this work were acquired from the text-mined dataset of 31,782 inorganic materials synthesis recipes originally extracted from the literature by Kononova et al.[33] and available at `https://github.com/CederGroupHub/text-mined-synthesis_public` (version 2020-07-13). The original dataset was filtered down to 8,530 reactions that contain: 1) precursors composed of ≤ 2 solids and ≤ 1 elemental gases (i.e., O_2, H_2, and N_2), 2) no elements with an atomic number greater than 94 or for which the Gibbs free energy descriptor does not apply (e.g., Ne, Ar, Pm, Ra), 3) ten or fewer total elements due to limitations in the convex hull algorithm, and 4) no ions. Finally, these reactions were required to be stoichiometrically balanceable after adjusting compositions for hydrates and fractional formulas. For reactions containing variable compositions with one open variable (e.g., $Nd_{1-x}Sr_xCoO_3$), we attempted to substitute all extracted values of x and retained the reactions that could be successfully balanced.

Competition metrics and free energies were assessed for each of the remaining reactions. For the enumerated competing reactions, metastable phases were considered up to a maximum threshold of $\Delta G_{hull} = 50$ meV/atom, evaluated at ambient temperature ($T = 300$ K). Interface reaction hulls were constructed at the maximum temperature reported during synthesis, T_{syn}. If this was not provided, a temperature of 800 ℃ was assumed. Formation energies, $\Delta G_f(T_{syn})$, were assigned based on the ground-state energy for a given composition; i.e., we selected the lowest available formation energy of all polymorphs with the composition of interest. For increased accuracy, we did not include a reaction if any of its entries were missing from our thermodynamic data. For reactions with an open gas (i.e., O_2, H_2, N_2), we assigned a chemical potential of $\mu_{gas} = 0$ eV (i.e., standard state at T_{syn}) for that element. For reactions completed in air, we assumed an O_2 partial pressure of 0.21 atm and thus assigned a chemical potential of $\mu_O = \frac{1}{2}k_b T_{syn} \ln(0.21)$ eV. Finally, we removed duplicates with the same reaction equation and temperature/environment, as well as identity reactions (e.g., $A \rightarrow A$), These filtering steps yielded a total of 3,520 unique literature reactions.

Precursor Materials

Precursors for all experiments were purchased from chemical providers or produced via known solid-state synthesis approaches, as necessary. Precursors acquired from chemical providers include: barium carbonate ($BaCO_3$, J.T.Baker 99.9%), titanium(IV) oxide (anatase TiO_2, Acros Organics 99.9%), barium sulfate ($BaSO_4$, J.T.Baker 99.9%), barium hydroxide hydrate ($Ba(OH)_2 \cdot 8\,H_2O$, Mathsen Colman & Bell 98%), barium chloride hydrate ($BaCl_2 \cdot 2\,H_2O$, Fisher Scientific 99.9%), activated carbon (C, J.T.Baker 99.9%), and titanium metal (Ti, annealed foil, Alfa Aesar 99.7%).

Ba_2TiO_4 was synthesized based on an existing approach[48] using stoichiometric amounts of $BaCO_3$ and anatase TiO_2. The chemicals were mixed, ground using a mortar and pestle, placed in an alumina boat inside of a mullite process tube with self-sealing endcaps, and then heated at 950 ℃ for 16 hrs under Ar flow with a heating rate of 10 ℃/min. The powder was then reground and reheated at 1100 ℃ for another 16 hrs at a heating rate of 10 ℃/min. Handling operations were completed in an Ar glovebox due to the hygroscopic nature of Ba_2TiO_4.

$BaTi_2O_5$ was synthesized with stoichiometric amounts of $BaCO_3$ and anatase TiO_2. The chemicals were mixed, ground using a mortar and pestle, and heated in an alumina boat at 900 ℃ for 5 hrs as a pre-treatment step. The powder was then reground and reheated at 1220-1225 ℃ for 24 hrs with heating and cooling steps of 3 hrs.[49]

Barium sulfide (BaS) was synthesized following an existing approach[50] using $BaSO_4$ and activated carbon. The chemicals were mixed, ground using a mortar and pestle, pressed into a 0.5” diameter pellet with two tons of force, and heated in an alumina boat at 1100 ℃ for 7-10 min in air, with a heating rate of 10 ℃/min and natural cooling in the furnace.

Sodium metatitanate (Na_2TiO_3) was synthesized based on an existing approach[51] using stoichiometric amounts of sodium hydroxide (NaOH, Fisher Scientific 99.9%) and anatase TiO_2, with a slight excess of NaOH. The chemicals were mixed, ground using a mortar and pestle, and heated in an alumina boat at 500 ℃ for 2 hrs with a heating rate of 10 ℃/min.

Ex Post Facto SPXRD Reactions

Synchrotron powder X-ray diffraction (SPXRD) data were collected in transmission (i.e., Debye-Scherrer) geometry on beamline 28-ID-2 (XPD) at the National Synchrotron Light Source-II (NSLS-II). Data were collected on a 2D area detector (PerkinElmer XRD 1621, 2048x2048 pixel array, 200x200 μm pixel size) at a sample-to-detector distance of 1407.1 mm using an incident X-ray energy of 68.12 keV ($\lambda = 0.182$ Å) with a 0.60 x 0.20 mm beam size. A total acquisition time of 1 s was used, summing five subframes collected for 0.2 s each.

Samples were packed into 1.1 mm OD/0.9 mm ID quartz capillaries. The capillary ends were filled with a 3 mm plug of powder silicon (Si, Strem 99.0%) followed by a cap of recycled silicon dioxide (SiO_2). To account for possible gas production, the capillaries for Expts. 1, 2, 5, 7, and 8 were left unsealed, and a moderate vacuum ($P_{gage} = -20$ in Hg) was pulled on the samples during heating. All other sample capillaries (Expts. 3, 4, 6, and 9) were flame-sealed under argon.

Experiments were carried out in the gradient furnace described in Ref. 43, which heats samples to different temperatures across a range of spatial positions on the capillary (Figure S19). The furnace was operated with a Eurotherm 2408 temperature controller and a TDK Lambda 900W (30V/30A) power supply. Furnace heating elements were wound from resistive wire (Kanthal A-1, #24 awg). A K-type thermocouple (stainless steel, 0.01" OD) placed at an intermediate position along the sample was used as input for PID control of the furnace. We performed experiments in three temperature ranges with setpoints of $T_H = 550$ ℃ (Expts. 1, 3, 4, 8), $T_{L1} = 450$ ℃ (Expts. 2, 5, 6, 9), and $T_{L2} = 400$ ℃ (Expt. 7). This choice was motivated by differences in reactivity among the samples. Position-dependent temperatures were determined using a fit of measured *in situ* lattice expansion from NaCl/Si and Al_2O_3/MgO standards (Figure S20). Using the root-mean-square error of the curve fit, the estimated uncertainty for each temperature point is 11.4 ℃ (T_H), 7.9 ℃ (T_{L1}), and 10.9 ℃ (T_{L2}). The experiments spanned a total temperature range of 189-1064 ℃.

The median total time of each experiment was ~67 min. Heating, holding, and cooling

times varied among experiments due to differences in sample heat capacities, reactivities, and the sizes of investigated temperature windows; specifically, unreactive samples (Expts. 8, 9) and samples with smaller studied temperature windows (Expts. 3, 4) were held at elevated temperatures for shorter times. These differences are accounted for in our analysis via the use of relative metrics (i.e., mole fraction) and normalization by reaction progress. The exact heating, holding, and cooling times for each sample are shown in Table S2.

Quantitative Phase Analysis of Powder Diffraction Data

Quantitative analysis of synchrotron powder X-ray powder diffraction data was carried out with the Rietveld method using either the TOPAS v6 (Expts. 1-3, 5, 7-9) or GSAS-II (Expts. 4, 6) software packages.[52,53] Atomic displacement parameters were fixed to $B = 1$ Å^2, and peak broadening was primarily modeled via crystal size broadening using a Lorentzian function. Site occupancies were fixed at one.

Supporting Information

- Description of cost function power transformation; interface reaction hull for BaO|TiO$_2$; open-O$_2$ reaction data for BaTiO$_3$; selected Rietveld refinements from all experiments; interface reaction hull plots for selected experimental reactions; selectivity metric correlations; details of carbonate correction; description of gradient furnace setup and temperature calibration; table of experimental reaction times and temperature setpoints (PDF)

- Table of 3,520 reactions extracted from the literature with their calculated performance metrics and DOIs (XLSX)

- Table of 82,985 predicted/ranked closed BaTiO$_3$ synthesis reactions (XLSX)

- Table of 62,133 predicted/ranked open-O$_2$ BaTiO$_3$ synthesis reactions (XLSX)

- Table of 133 Pareto optimal open-O_2 $BaTiO_3$ synthesis reactions (XLSX)

Acknowledgements

Synopsis

Selectivity analysis of massive inorganic chemical reaction networks identifies alternative synthesis recipes for solid-state materials (e.g., $BaTiO_3$) that outperform conventional approaches.

References

(1) Kovnir, K. Predictive Synthesis. *Chemistry of Materials* **2021**, *33*, 4835–4841, DOI: `10.1021/acs.chemmater.1c01484`.

(2) DiSalvo, F. J. Solid-State Chemistry: A Rediscovered Chemical Frontier. *Science (New York, N.Y.)* **1990**, *247*, 649–655, DOI: `10.1126/science.247.4943.649`.

(3) Kohlmann, H. Looking into the Black Box of Solid-State Synthesis. *European Journal of Inorganic Chemistry* **2019**, *2019*, 4174–4180, DOI: `10.1002/ejic.201900733`.

(4) Corey, E.; Cheng, X.-M. *The Logic of Chemical Synthesis*; John Wiley & Sons, Inc., 1989.

(5) Coley, C. W.; Green, W. H.; Jensen, K. F. Machine Learning in Computer-Aided Synthesis Planning. *Accounts of Chemical Research* **2018**, *51*, 1281–1289, DOI: `10.1021/acs.accounts.8b00087`.

(6) 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, DOI: `10.1021/acs.chemmater.7b03500`.

(7) Narayan, A.; Bhutani, A.; Rubeck, S.; Eckstein, J. N.; Shoemaker, D. P.; Wagner, L. K. Computational and Experimental Investigation for New Transition Metal Selenides and Sulfides: The Importance of Experimental Verification for Stability. *Physical Review B* **2016**, *94*, 045105, DOI: `10.1103/PhysRevB.94.045105`.

(8) Rong, Z.; Xiao, P.; Liu, M.; Huang, W.; Hannah, D. C.; Scullin, W.; Persson, K. A.; Ceder, G. Fast Mg^{2+} Diffusion in $Mo_3(PO_4)_3O$ for Mg Batteries. *Chemical Communications* **2017**, *53*, 7998–8001, DOI: `10.1039/c7cc02903a`.

(9) Gardner, J.; Thakre, A.; Kumar, A.; Scott, J. F. Tin Titanate—the Hunt for a

New Ferroelectric Perovskite. *Reports on Progress in Physics* **2019**, *82*, 092501, DOI: `10.1088/1361-6633/ab37d4`.

(10) West, A. R. *Solid State Chemistry and Its Applications*, 2nd ed.; Wiley: Chichester, West Sussex, UK, 2014.

(11) 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. Observing and Modeling the Sequential Pairwise Reactions That Drive Solid-State Ceramic Synthesis. *Advanced Materials* **2021**, *33*, 2100312, DOI: `10.1002/adma.202100312`.

(12) Richards, W. D.; Miara, L. J.; Wang, Y.; Kim, J. C.; Ceder, G. Interface Stability in Solid-State Batteries. *Chemistry of Materials* **2015**, *28*, 266–273, DOI: `10.1021/acs.chemmater.5b04082`.

(13) 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*, 1088–1095, DOI: `10.1038/s41563-020-0688-6`.

(14) Aykol, M.; Montoya, J. H.; Hummelshøj, J. Rational Solid-State Synthesis Routes for Inorganic Materials. *Journal of the American Chemical Society* **2021**, *143*, 9244–9259, DOI: `10.1021/jacs.1c04888`.

(15) d'Heurle, F. M. Nucleation of a New Phase from the Interaction of Two Adjacent Phases: Some Silicides. *Journal of Materials Research* **1988**, *3*, 167–195, DOI: `10.1557/jmr.1988.0167`.

(16) Kumar, M. M.; Palkar, V. R.; Srinivas, K.; Suryanarayana, S. V. Ferroelectricity in a Pure $BiFeO_3$ Ceramic. *Applied Physics Letters* **2000**, *76*, 2764–2766, DOI: `10.1063/1.126468`.

(17) Bernardo, M.; Jardiel, T.; Peiteado, M.; Caballero, A.; Villegas, M. Reaction Pathways in the Solid State Synthesis of Multiferroic $BiFeO_3$. *Journal of the European Ceramic Society* **2011**, *31*, 3047–3053, DOI: `10.1016/j.jeurceramsoc.2011.03.018`.

(18) Merkle, R.; Maier, J. On the Tammann–Rule. *Zeitschrift für anorganische und allgemeine Chemie* **2005**, *631*, 1163–1166, DOI: `10.1002/zaac.200400540`.

(19) Huo, H.; Bartel, C. J.; He, T.; Trewartha, A.; Dunn, A.; Ouyang, B.; Jain, A.; Ceder, G. Machine-Learning Rationalization and Prediction of Solid-State Synthesis Conditions. *Chemistry of Materials* **2022**, *34*, 7323–7336, DOI: `10.1021/acs.chemmater.2c01293`.

(20) Jia, X.; Lynch, A.; Huang, Y.; Danielson, M.; Lang'at, I.; Milder, A.; Ruby, A. E.; Wang, H.; Friedler, S. A.; Norquist, A. J.; Schrier, J. Anthropogenic Biases in Chemical Reaction Data Hinder Exploratory Inorganic Synthesis. *Nature* **2019**, *573*, 251–255, DOI: `10.1038/s41586-019-1540-5`.

(21) He, T.; Huo, H.; Bartel, C. J.; Wang, Z.; Cruse, K.; Ceder, G. Precursor Recommendation for Inorganic Synthesis by Machine Learning Materials Similarity from Scientific Literature. *Science Advances* **2023**, *9*, eadg8180, DOI: `10.1126/sciadv.adg8180`.

(22) Jiang, Z.; Arun Ramanathan,; Shoemaker, D. P. In Situ Identification of Kinetic Factors That Expedite Inorganic Crystal Formation and Discovery. *Journal of Materials Chemistry C* **2017**, *5*, 5709–5717, DOI: `10.1039/C6TC04931A`.

(23) Miura, A.; Ito, H.; Bartel, C. J.; Sun, W.; Rosero-Navarro, N. C.; Tadanaga, K.; Nakata, H.; Maeda, K.; Ceder, G. Selective Metathesis Synthesis of $MgCr_2S_4$ by Control of Thermodynamic Driving Forces. *Materials Horizons* **2020**, *7*, 1310–1316, DOI: `10.1039/c9mh01999e`.

(24) Todd, P. K.; McDermott, M. J.; Rom, C. L.; Corrao, A. A.; Denney, J. J.; Dwaraknath, S. S.; Khalifah, P. G.; Persson, K. A.; Neilson, J. R. Selectivity in Yttrium

Manganese Oxide Synthesis via Local Chemical Potentials in Hyperdimensional Phase Space. *Journal of the American Chemical Society* **2021**, *143*, 15185–15194, DOI: `10.1021/jacs.1c06229`.

(25) Chen, J.; Cross, S. R.; Miara, L. J.; Cho, J.-J.; Wang, Y.; Sun, W. Navigating Phase Diagram Complexity to Guide Robotic Inorganic Materials Synthesis. 2023; `http://arxiv.org/abs/2304.00743`.

(26) McDermott, M. J.; Dwaraknath, S. S.; Persson, K. A. A Graph-Based Network for Predicting Chemical Reaction Pathways in Solid-State Materials Synthesis. *Nature Communications* **2021**, *12*, 3097, DOI: `10.1038/s41467-021-23339-x`.

(27) 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, DOI: `10.1063/1.4812323`.

(28) 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, DOI: `10.1038/s41467-018-06682-4`.

(29) Chase, J. M. W. *NIST-JANAF Thermochemical Tables*, 4th ed.; American Institute of Physics for the National Institute of Standards and Technology: Washington D.C., 1998.

(30) Schmalzried, H. *Solid State Reactions*, 2nd ed.; Verlag Chemie, Weinheim: Germany, 1981.

(31) van Loo, F. J. J. Multiphase Diffusion in Binary and Ternary Solid-

State Systems. *Progress in Solid State Chemistry* **1990**, *20*, 47–99, DOI: `10.1016/0079-6786(90)90007-3`.

(32) Stanley, R. P.; Fomin, S. Enumerative Combinatorics. **1999**, 159–285, DOI: `10.1017/cbo9780511609589.005`.

(33) 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, DOI: `10.1038/s41597-019-0224-1`.

(34) Cooper, J. K.; Scott, S. B.; Ling, Y.; Yang, J.; Hao, S.; Li, Y.; Toma, F. M.; Stutzmann, M.; Lakshmi, K. V.; Sharp, I. D. Role of Hydrogen in Defining the N-Type Character of $BiVO_4$ Photoanodes. *Chemistry of Materials* **2016**, *28*, 5761–5771, DOI: `10.1021/acs.chemmater.6b01994`.

(35) Selbach, S. M.; Tybell, T.; Einarsrud, M.-A.; Grande, T. High-Temperature Semiconducting Cubic Phase of $BiFe_{0.7}Mn_{0.3}O_{3+\delta}$. *Physical Review B* **2009**, *79*, 214113, DOI: `10.1103/PhysRevB.79.214113`.

(36) Forrester, J. S.; Zobec, J. S.; Phelan, D.; Kisi, E. H. Synthesis of $PbTiO_3$ Ceramics Using Mechanical Alloying and Solid State Sintering. *Journal of Solid State Chemistry* **2004**, *177*, 3553–3559, DOI: `10.1016/j.jssc.2004.06.005`.

(37) Paik, Y.; Grey, C. P.; Johnson, C. S.; Kim, J.-S.; Thackeray, M. M. Lithium and Deuterium NMR Studies of Acid-Leached Layered Lithium Manganese Oxides. *Chemistry of Materials* **2002**, *14*, 5109–5115, DOI: `10.1021/cm0206385`.

(38) Herstedt, M.; Stjerndahl, M.; Nytén, A.; Gustafsson, T.; Rensmo, H.; Siegbahn, H.; Ravet, N.; Armand, M.; Thomas, J. O.; Edström, K. Surface Chemistry of Carbon-Treated $LiFePO_4$ Particles for Li-Ion Battery Cathodes Studied by PES. *Electrochemical and Solid-State Letters* **2003**, *6*, A202, DOI: `10.1149/1.1594413`.

(39) Kitamura, N.; Iwatsuki, H.; Idemoto, Y. Improvement of Cathode Performance of LiMn$_2$O$_4$ as a Cathode Active Material for Li Ion Battery by Step-by-Step Supersonic-Wave Treatments. *Journal of Power Sources* **2009**, *189*, 114–120, DOI: `10.1016/j.jpowsour.2008.10.046`.

(40) Neilson, J. R.; McDermott, M. J.; Persson, K. A. Modernist Materials Synthesis: Finding Thermodynamic Shortcuts with Hyperdimensional Chemistry. *Journal of Materials Research* **2023**, *38*, 2885–2893, DOI: `10.1557/s43578-023-01037-2`.

(41) Bierach, J. W. The Formation of Barium Titanate Ceramics by Solid State Reaction. Ph.D. thesis, Lawrence Berkeley National Laboratory, 1988.

(42) Hong, Q.-J.; Ushakov, S. V.; van de Walle, A.; Navrotsky, A. Melting Temperature Prediction Using a Graph Neural Network Model: From Ancient Minerals to New Materials. *Proceedings of the National Academy of Sciences* **2022**, *119*, e2209630119, DOI: `10.1073/pnas.2209630119`.

(43) O'Nolan, D.; Huang, G.; Kamm, G. E.; Grenier, A.; Liu, C.-H.; Todd, P. K.; Wustrow, A.; Tran, G. T.; Montiel, D.; Neilson, J. R.; Billinge, S. J. L.; Chupas, P. J.; Thornton, K. S.; Chapman, K. W. A Thermal-Gradient Approach to Variable-Temperature Measurements Resolved in Space. *Journal of Applied Crystallography* **2020**, *53*, 662–670, DOI: `10.1107/s160057672000415x`.

(44) 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, DOI: `10.1021/acs.chemmater.6b04861`.

(45) Szymanski, N. J.; Zeng, Y.; Huo, H.; Bartel, C. J.; Kim, H.; Ceder, G. Toward Autonomous Design and Synthesis of Novel Inorganic Materials. *Materials Horizons* **2021**, *8*, 2169–2198, DOI: `10.1039/d1mh00495f`.

(46) Ganose, A.; Jain, A.; Rignanese, G.-M.; Waroquiers, D.; Petretto, G., et al. Jobflow. GitHub, 2023; `https://github.com/materialsproject/jobflow`.

(47) Jain, A.; Ong, S. P.; Chen, W.; Medasani, B.; Qu, X.; Kocher, M.; Brafman, M.; Petretto, G.; Rignanese, G.; Hautier, G.; Gunter, D.; Persson, K. A. FireWorks: A Dynamic Workflow System Designed for High-throughput Applications. *Concurrency and Computation: Practice and Experience* **2015**, *27*, 5037–5059, DOI: `10.1002/cpe.3505`.

(48) McSloy, A. J.; Trussov, I.; Jarvis, A.; Cooke, D. J.; Slater, P. R.; Panchmatia, P. M. Mechanism of Carbon Dioxide and Water Incorporation in Ba_2TiO_4: A Joint Computational and Experimental Study. *The Journal of Physical Chemistry C* **2018**, *122*, 1061–1069, DOI: `10.1021/acs.jpcc.7b10330`.

(49) Zhu, N.; West, A. R. Formation and Stability of Ferroelectric $BaTi_2O_5$. *Journal of the American Ceramic Society* **2010**, *93*, 295–300, DOI: `10.1111/j.1551-2916.2009.03393.x`.

(50) Murthy, J. S. N.; Reddy, P. V. V. Solid-State Reaction Between Barium Sulfate and Carbon. *Chemical Engineering Communications* **2012**, *199*, 966–990, DOI: `10.1080/00986445.2011.631204`.

(51) Meng, F.; Liu, Y.; Xue, T.; Su, Q.; Wang, W.; Qi, T. Structures, Formation Mechanisms, and Ion-Exchange Properties of α-, β-, and γ-Na_2TiO_3. *RSC Advances* **2016**, *6*, 112625–112633, DOI: `10.1039/C6RA16984H`.

(52) Coelho, A. A. TOPAS and TOPAS-Academic: An Optimization Program Integrating Computer Algebra and Crystallographic Objects Written in C++. *Journal of Applied Crystallography* **2018**, *51*, 210–218, DOI: `10.1107/S1600576718000183`.

(53) Toby, B.; Dreele, R. V. GSAS-II: The Genesis of a Modern Open-Source All Purpose Crystallography Software Package. *Journal of Applied Crystallography* **2013**, *46*, 544–549, DOI: `10.1107/s0021889813003531`.

TOC Graphic
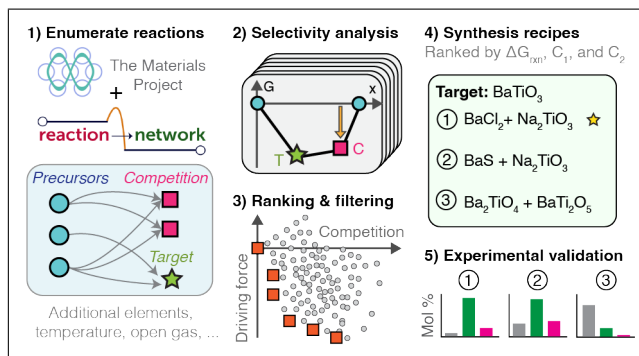
Supporting Information:

Assessing Thermodynamic Selectivity of Solid-State Reactions for the Predictive Synthesis of Inorganic Materials

Matthew J. McDermott,[†,‡,@] Brennan C. McBride,[¶,@] Corlyn Regier,[¶] Gia Thinh Tran,[¶] Yu Chen,[†,‡] Adam A. Corrao,[§] Max C. Gallant,[†,‡] Gabrielle E. Kamm,[§] Christopher J. Bartel,[‖] Karena W. Chapman,[§] Peter G. Khalifah,[§,⊥] Gerbrand Ceder,[†,‡] James R. Neilson,[¶] and Kristin A. Persson[*,#,‡]

†Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA
‡Department of Materials Science and Engineering, University of California, Berkeley, CA
¶Department of Chemistry, Colorado State University, Fort Collins, CO
§Department of Chemistry, Stony Brook University, Stony Brook, NY
‖Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN
⊥Chemistry Division, Brookhaven National Laboratory, Upton, NY
#Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA
@These authors contributed equally to this work.

E-mail: kapersson@lbl.gov

Power transform of costs in literature reactions
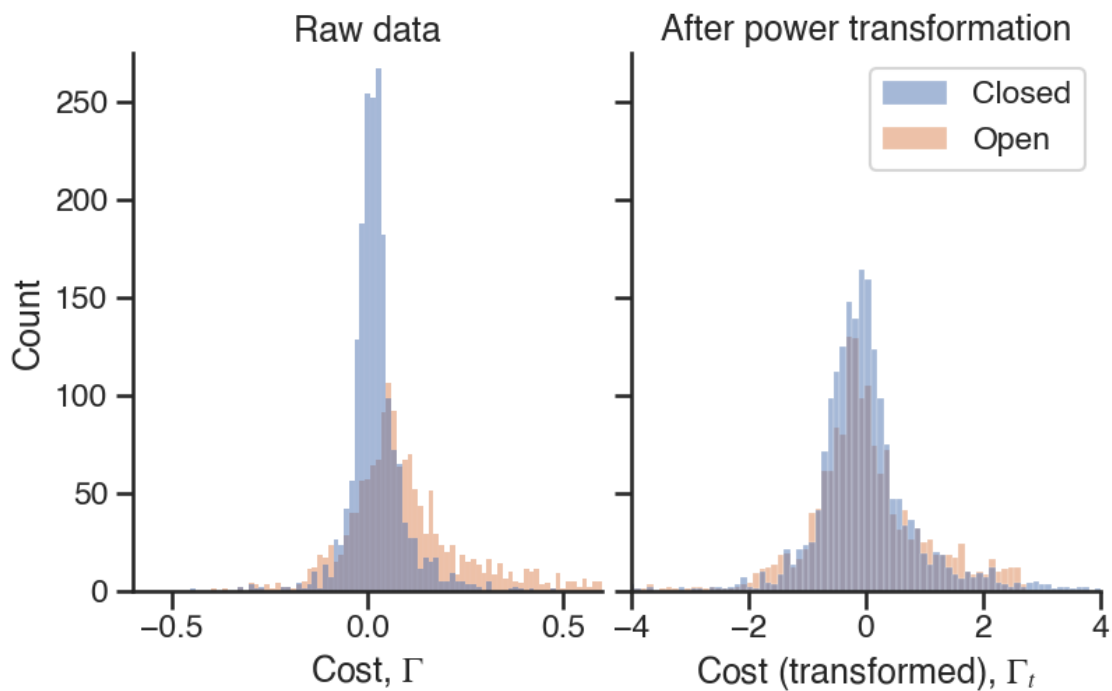
Figure S1: Power transformation applied to literature reaction costs, Γ. The power transform monotonically transforms the reaction costs for closed and open reactions such that they resemble standard normal distributions, allowing for better comparison between the two datasets.

BaO|TiO$_2$ interface reaction hull
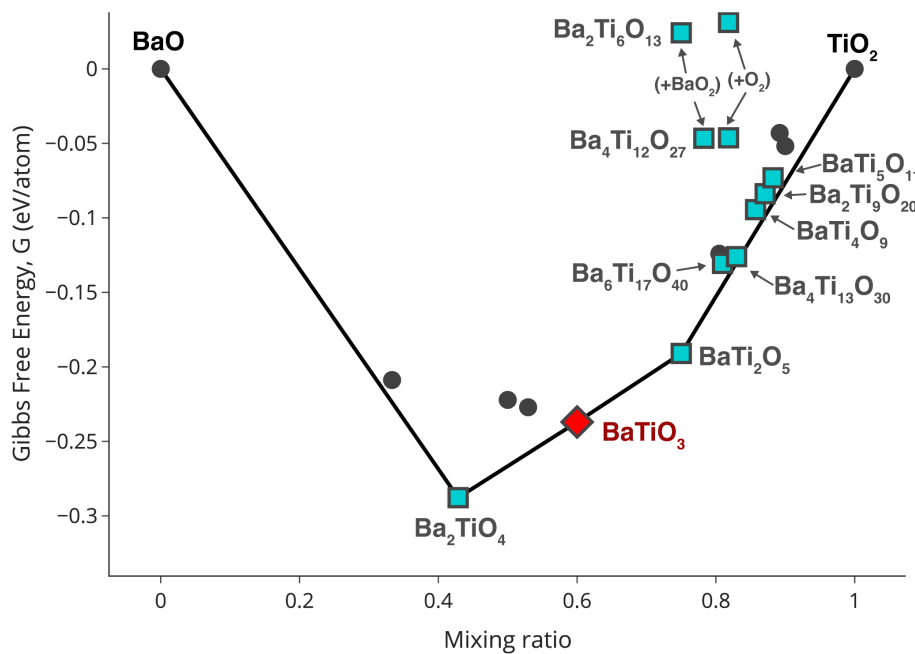
Figure S2: Calculated interface reaction hull for BaO|TiO$_2$ at $T = 600$ °C. Blue squares denote experimentally known ternary phases in the system. Ba$_2$TiO$_4$ is the most commonly observed intermediate/impurity in standard syntheses of BaTiO$_3$ from BaO (or BaCO$_3$) and TiO$_2$; it is also predicted to be the phase with the greatest driving force to form.

BaTiO$_3$ synthesis reactions: open O$_2$

Table S1: Selected experimental reactions to BaTiO$_3$ and their associated grand potential energies ($\mu_O = -0.1001$ eV), $\Delta\Phi_{\text{rxn}}$ ($T = 600\,$°C), primary competition scores, C_1, secondary competition scores, C_2, and costs, Γ.

Expt.	Reaction	$\Delta\Phi_{\text{rxn}}$ (eV/at)	C_1 (eV/at)	C_2 (eV/at)	Γ (eV/at)
1	BaCO$_3$ + TiO$_2$ $\longrightarrow$ BaTiO$_3$ + CO$_2$	0.112	0.114	0.001	0.063
2	BaO$_2$ + TiO$_2$ $\longrightarrow$ BaTiO$_3$ + 0.5 O$_2$	-0.590	0.078	0.389	0.151
3	Ba$_2$TiO$_4$ + TiO$_2$ $\longrightarrow$ 2 BaTiO$_3$	-0.089	0.085	0.114	0.081
4	Ba$_2$TiO$_4$ + BaTi$_2$O$_5$ $\longrightarrow$ 3 BaTiO$_3$	-0.002	-0.002	0.000	-0.001
5	Ba(OH)$_2$·H$_2$O + 3.666 Ti $\longrightarrow$ BaTiO$_3$ + 1.333 Ti$_2$H$_3$	-0.713	8.188	8.973	7.651
6	BaCl$_2$ + Na$_2$TiO$_3$ $\longrightarrow$ BaTiO$_3$ + 2 NaCl	-0.112	-0.004	0.077	0.021
7	BaS + Na$_2$TiO$_3$ $\longrightarrow$ BaTiO$_3$ + Na$_2$S	-0.077	4.032	4.174	3.685
8	2 BaS + 3 TiO$_2$ $\longrightarrow$ 2 BaTiO$_3$ + TiS$_2$	0.168	4.276	4.109	3.790
9	BaSO$_4$ + 2 TiO$_2$ $\longrightarrow$ BaTiO$_3$ + TiOSO$_4$	0.533	0.533	-0.000	0.293
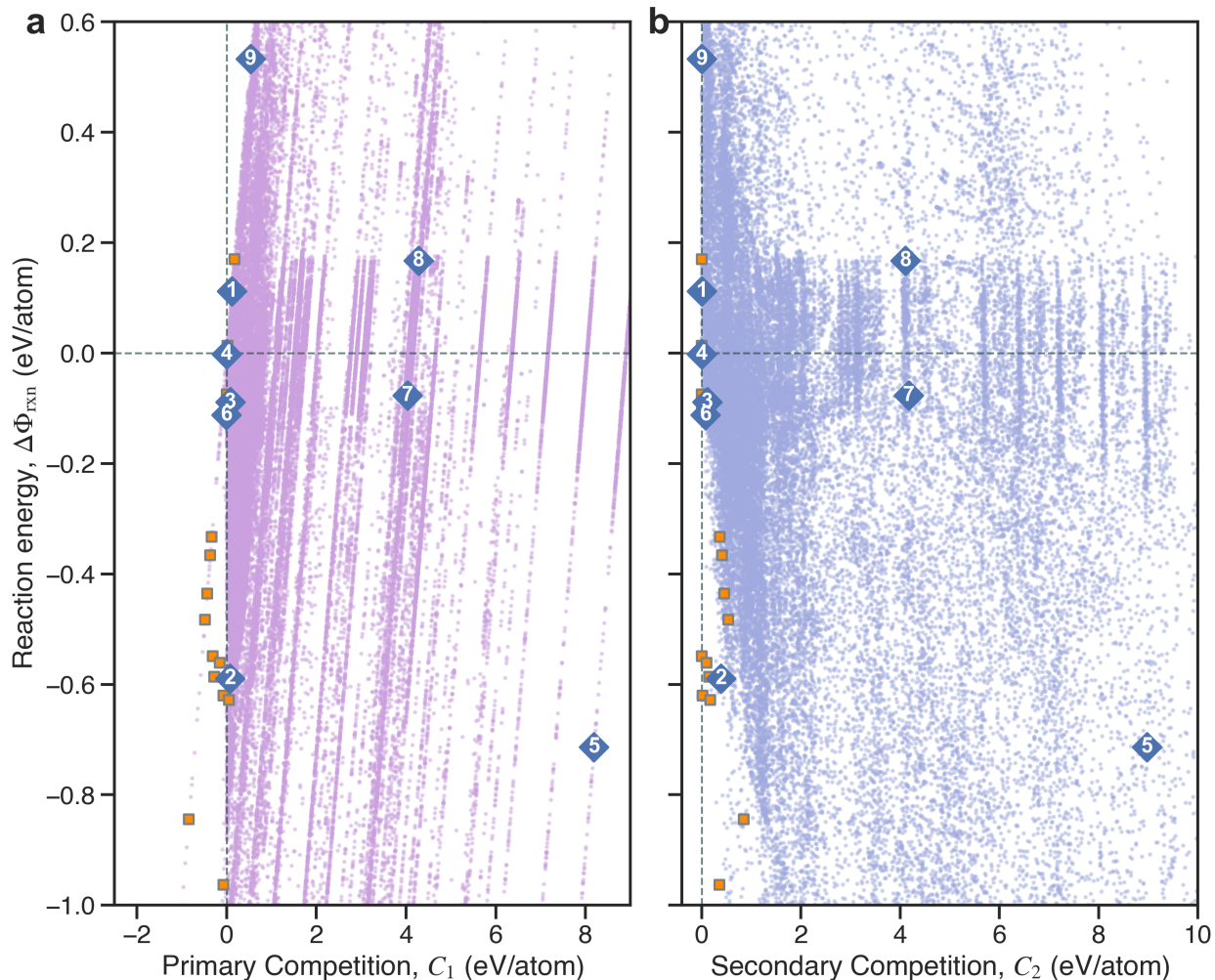
Figure S3: Synthesis map of 62,133 calculated open-O_2 synthesis reactions producing BaTiO$_3$. Reaction energies and competition scores are calculated assuming a temperature of $T = 600$ °C and oxygen chemical potential $\mu_O = -0.1001$ eV, as approximated by selected experimental conditions (vacuum at $P_{\text{gage}} = -20$ in Hg). As in Figure 5, reactions are plotted on a shared axis of reaction energy, $\Delta\Phi_{\text{rxn}}$, and on independent axes of (a) primary competition, C_1, and (b) secondary competition, C_2. Orange squares represent reactions on the three-dimensional Pareto frontier of $\Delta\Phi_{\text{rxn}}$, C_1, and C_2. Blue diamonds indicate selected reactions experimentally tested in this work.

Selected Rietveld refinements for BaTiO$_3$ experiments
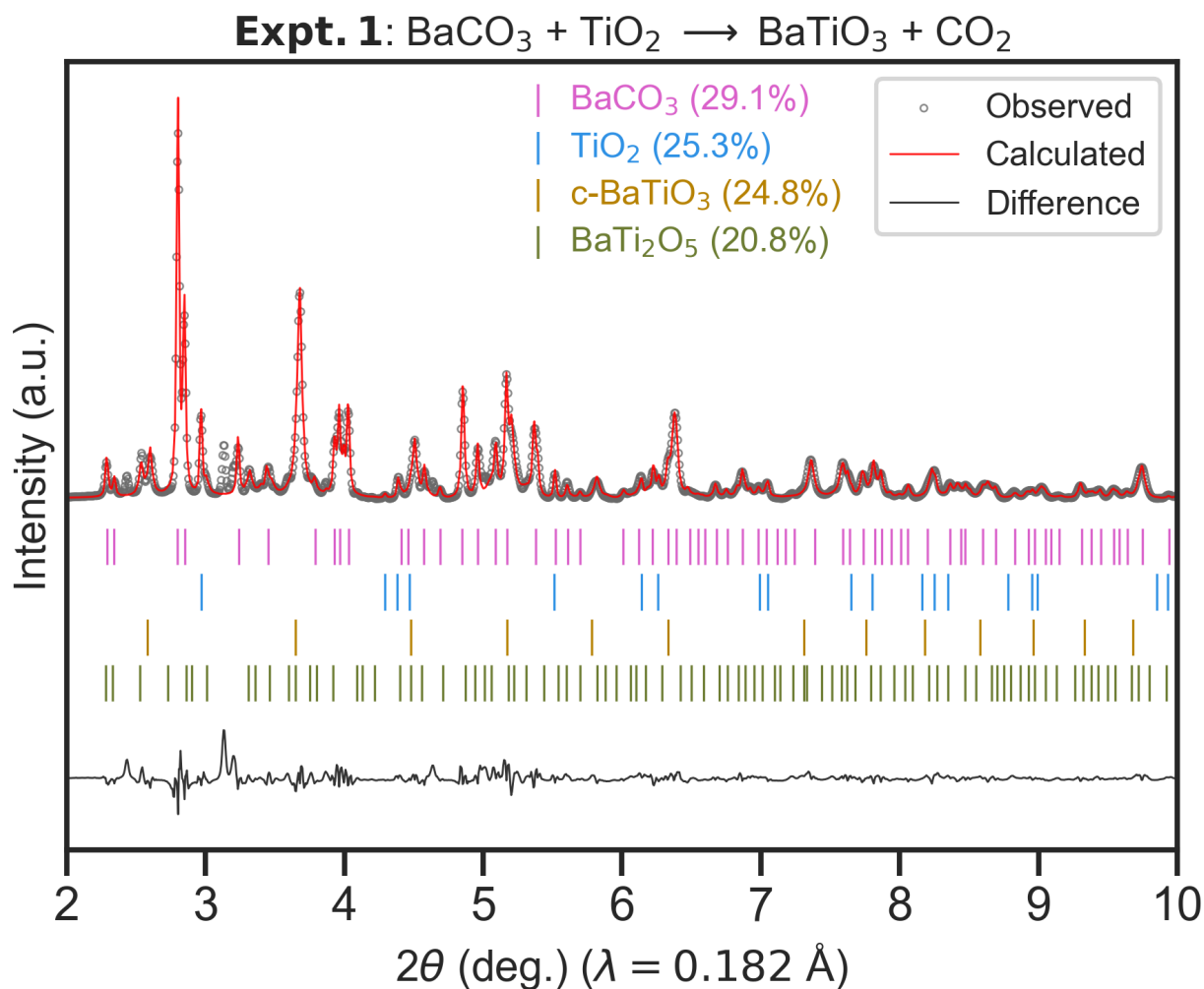
Figure S4: Selected Rietveld refinement for Experiment 1. The observed pattern represents *ex post facto* synchrotron powder X-ray diffraction data (SPXRD) captured following reaction at $T = 1040$ ℃, which corresponds to the temperature with the highest BaTiO$_3$ yield. Phase fractions are shown in units of mole percent.
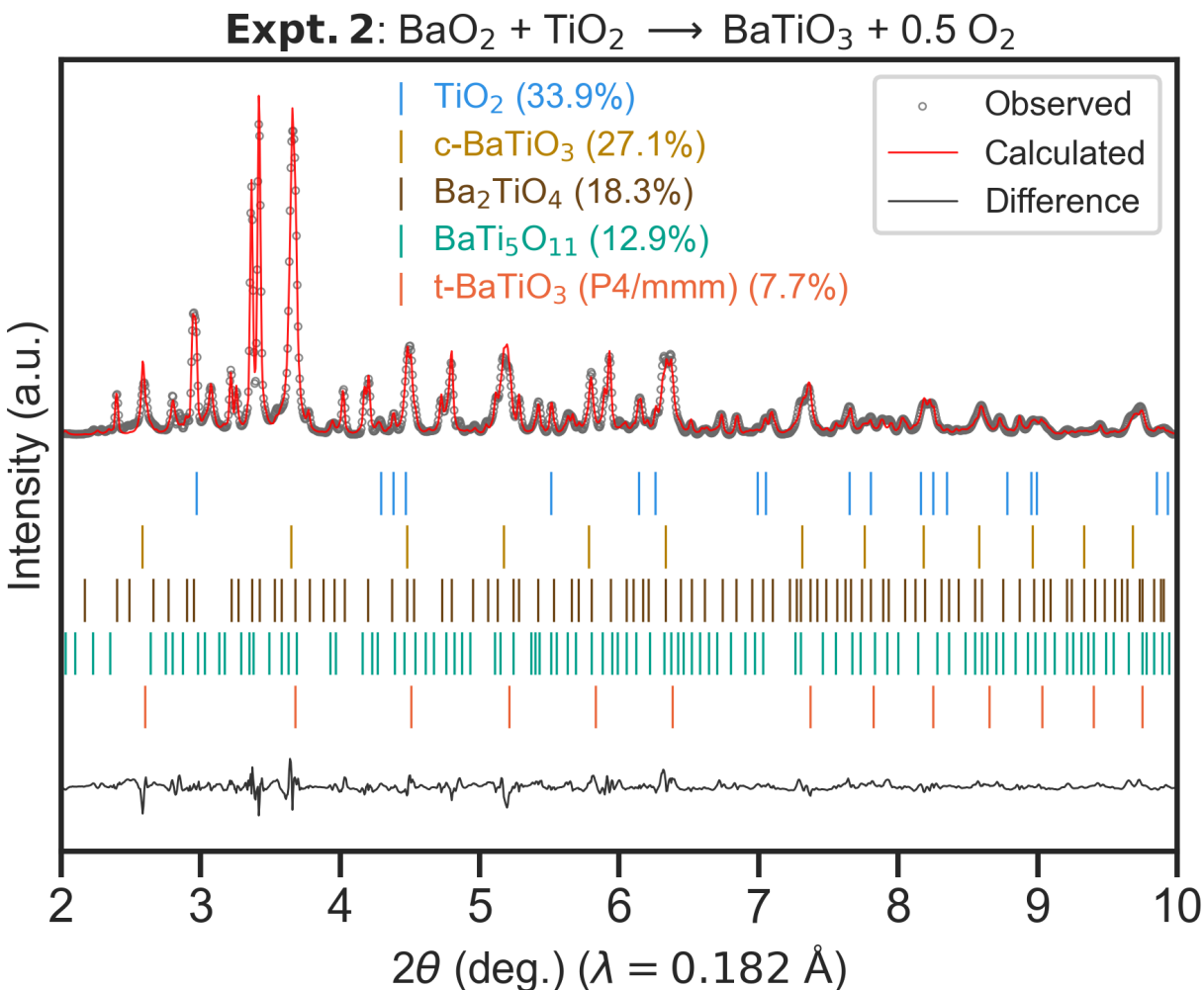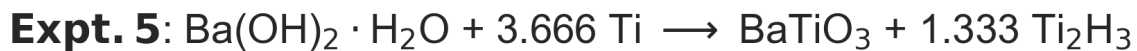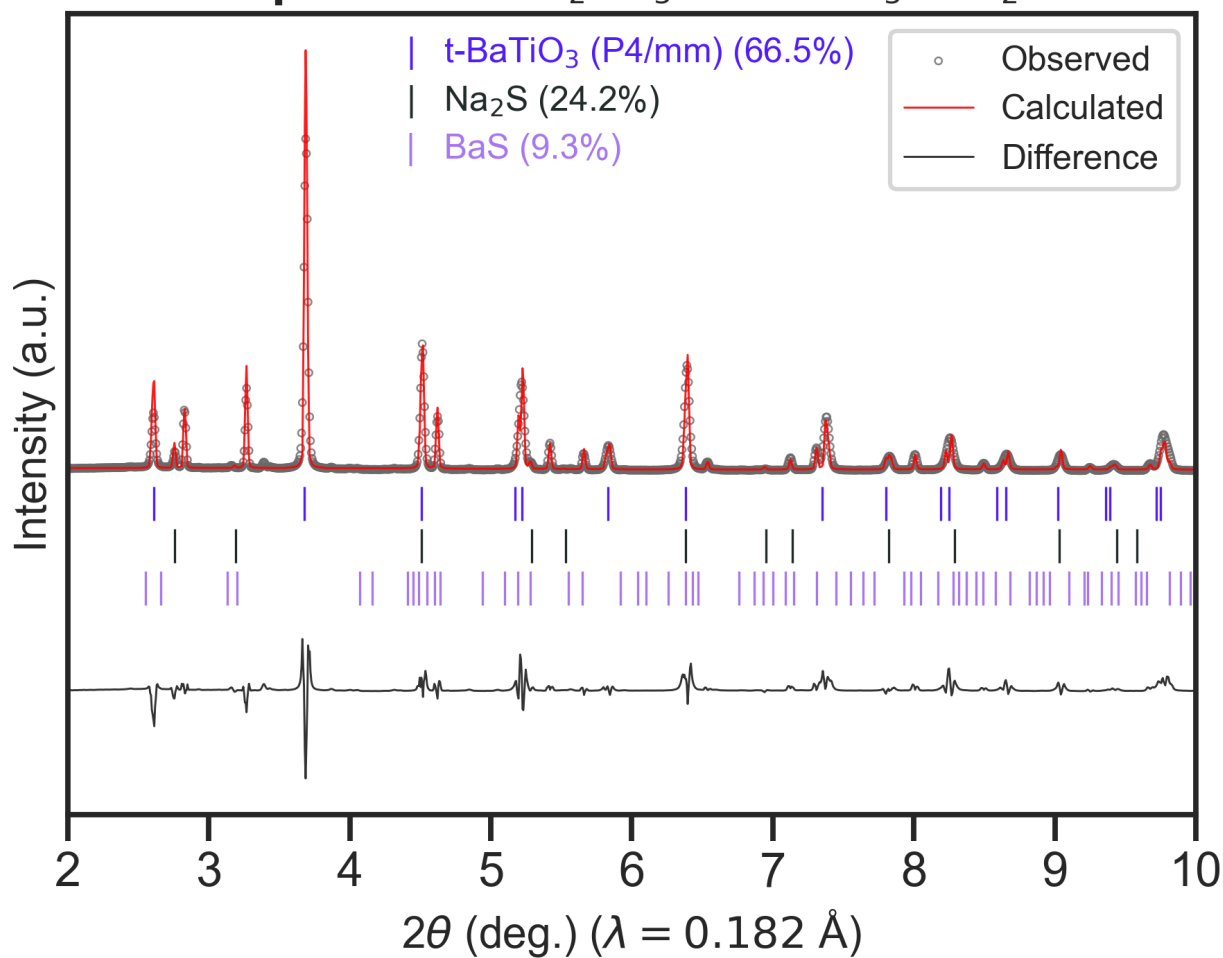
Figure S5: Selected Rietveld refinement for Experiment 2. The observed pattern represents *ex post facto* SPXRD data captured following reaction at $T = 762$ ℃, which corresponds to the temperature with the highest $BaTiO_3$ yield. Phase fractions are shown in units of mole percent.
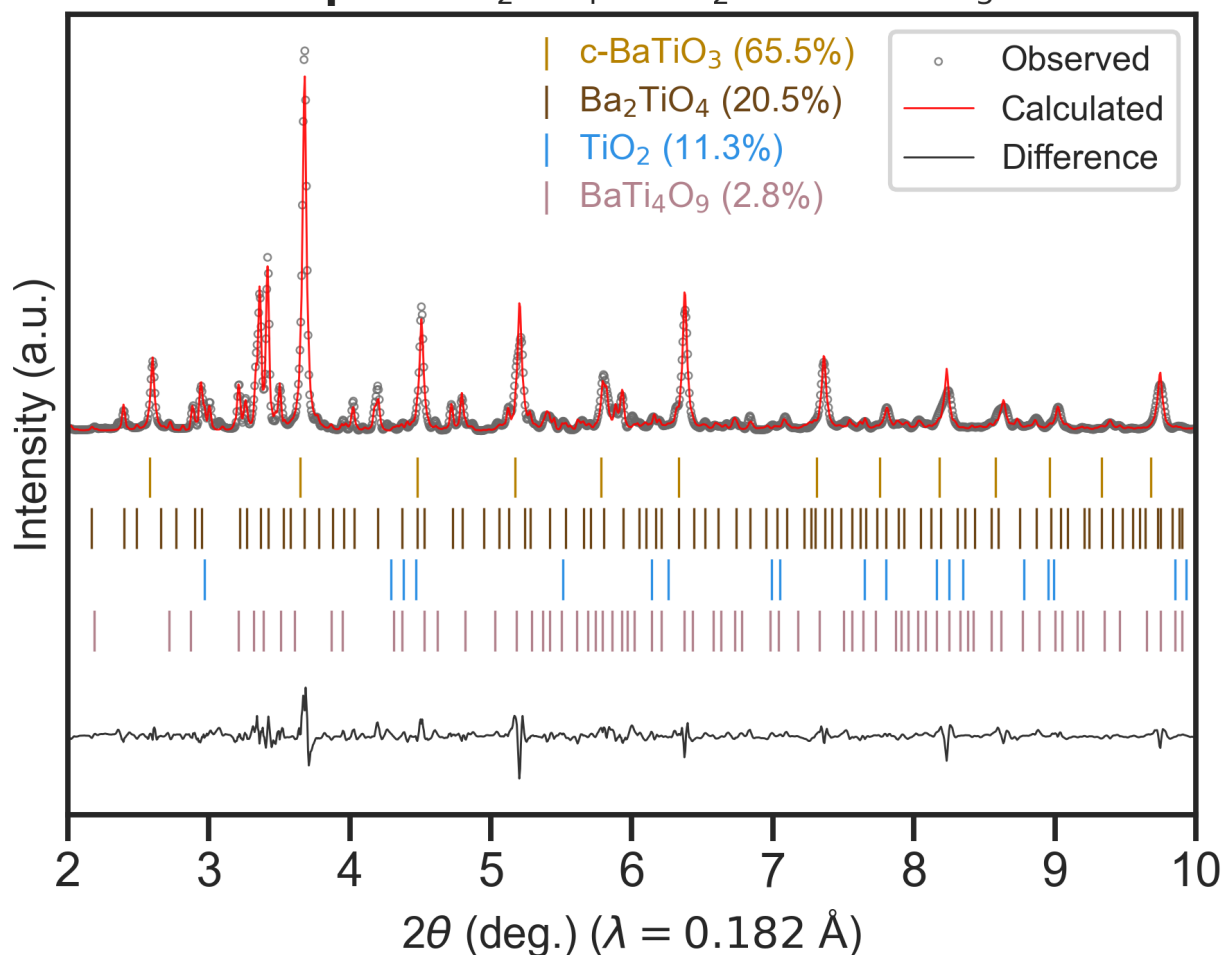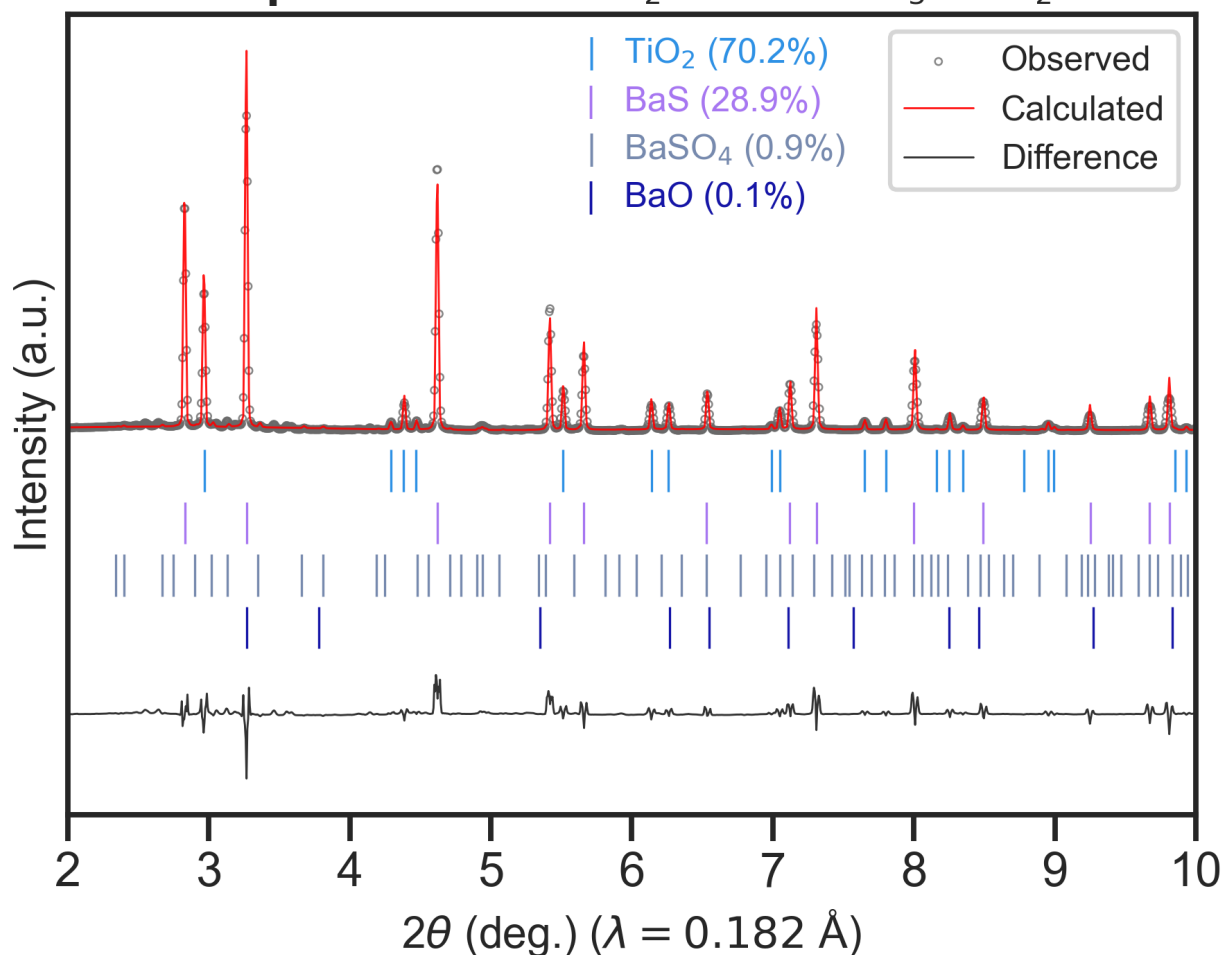
Figure S6: Selected Rietveld refinement for Experiment 3. The observed pattern represents *ex post facto* SPXRD data captured following reaction at $T = 1045$ ℃, which corresponds to the temperature with the highest $BaTiO_3$ yield. Phase fractions are shown in units of mole percent.
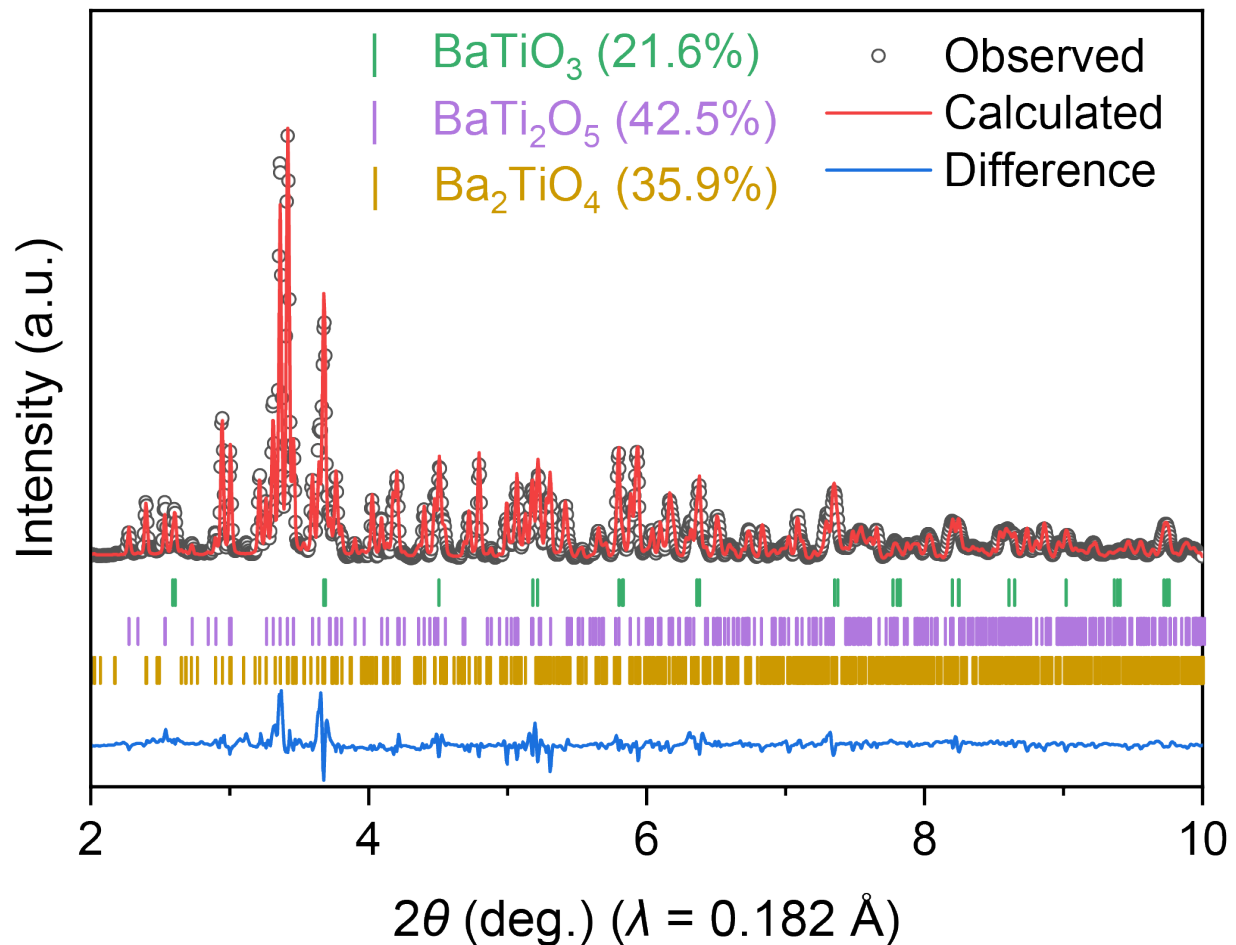
Figure S7: Selected Rietveld refinement for Experiment 4. The observed pattern represents *ex post facto* SPXRD data captured following reaction at $T = 1025$ ℃, which corresponds to the temperature with the highest $BaTiO_3$ yield. Phase fractions are shown in units of mole percent.
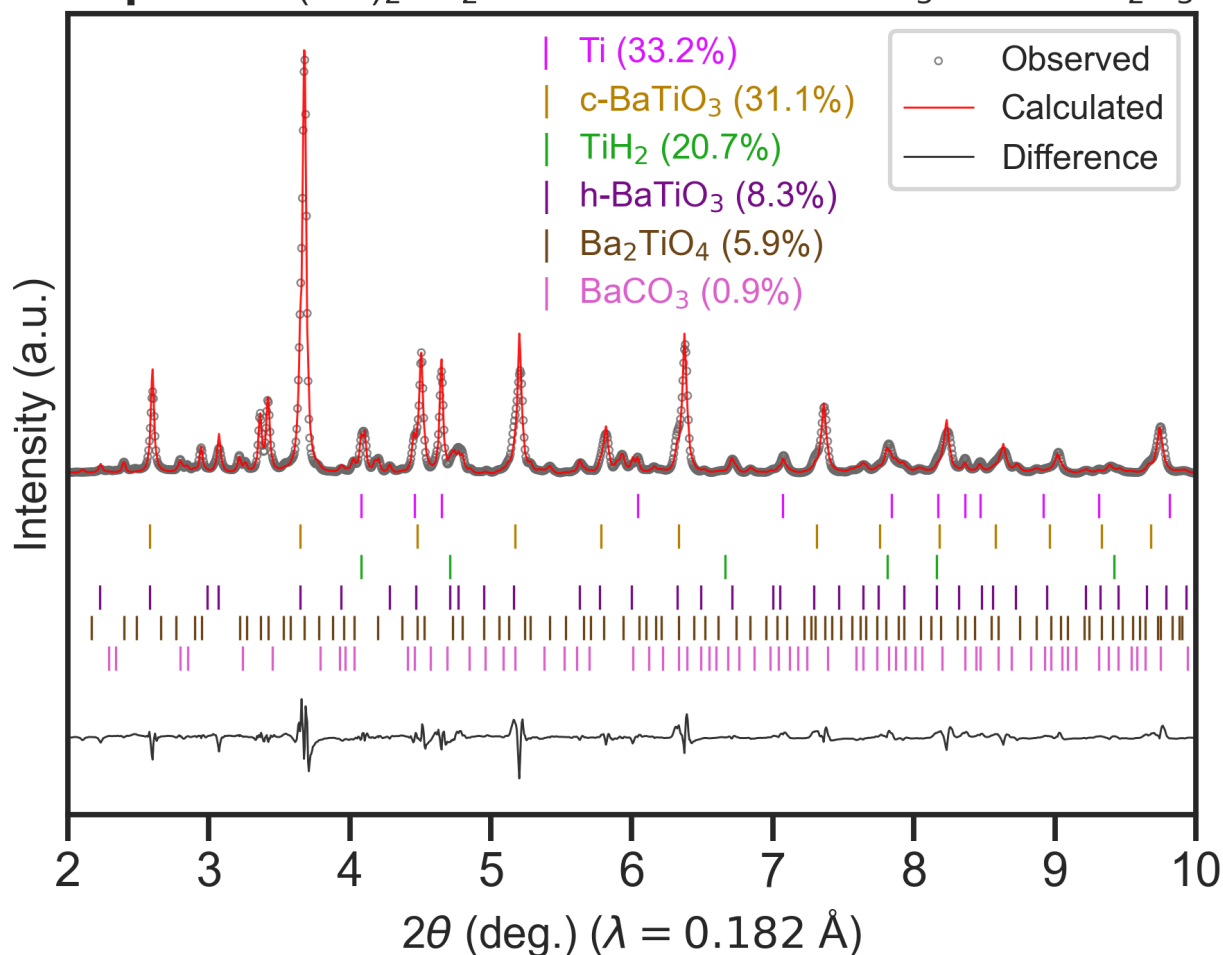
Figure S8: Selected Rietveld refinement for Experiment 5. The observed pattern represents *ex post facto* SPXRD data captured following reaction at $T = 474$ ℃, which corresponds to the temperature with the highest $BaTiO_3$ yield. Phase fractions are shown in units of mole percent.
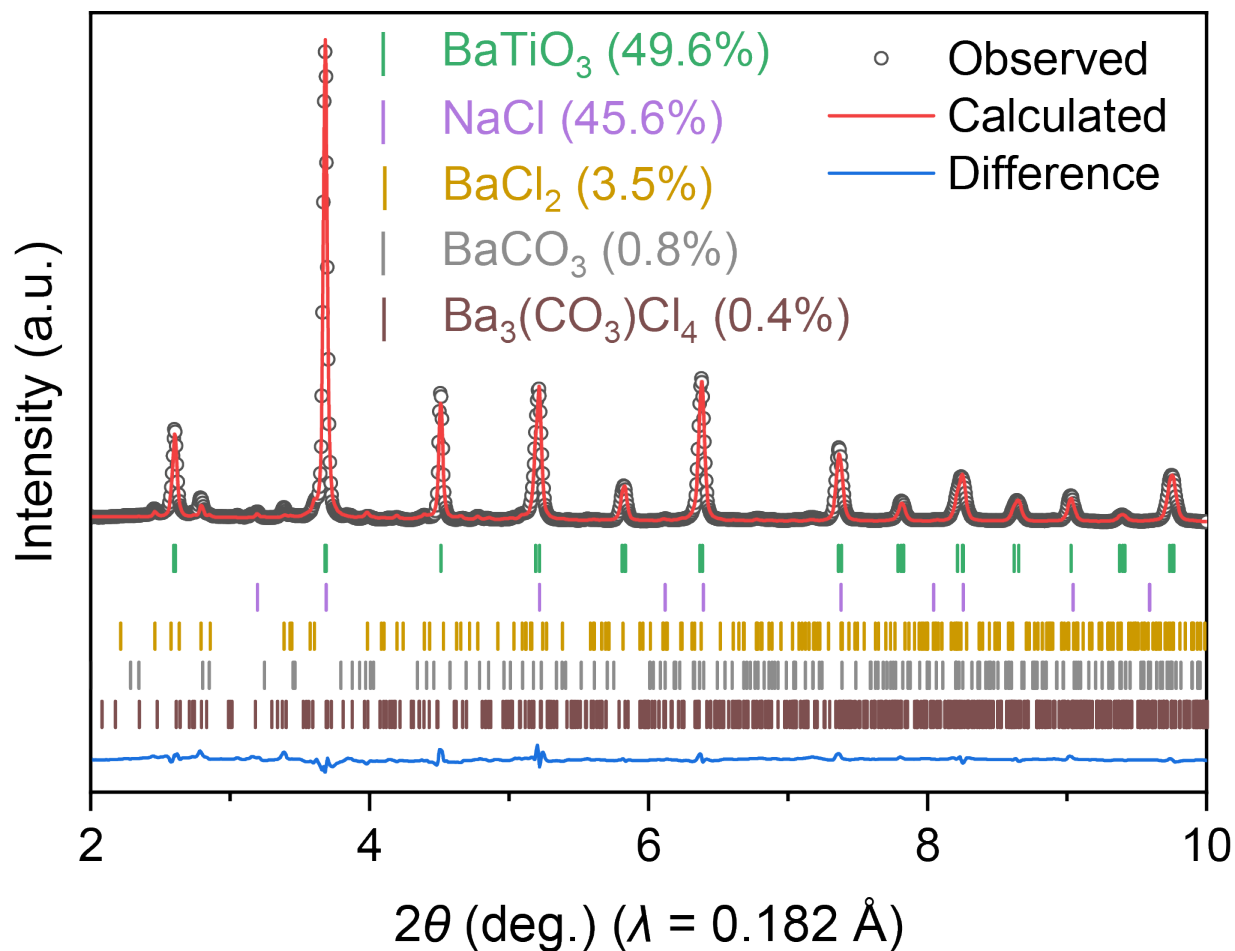
Figure S9: Selected Rietveld refinement for Experiment 6. The observed pattern represents *ex post facto* SPXRD data captured following reaction at $T = 625$ °C, which corresponds to the temperature with the highest $BaTiO_3$ yield. Phase fractions are shown in units of mole percent.
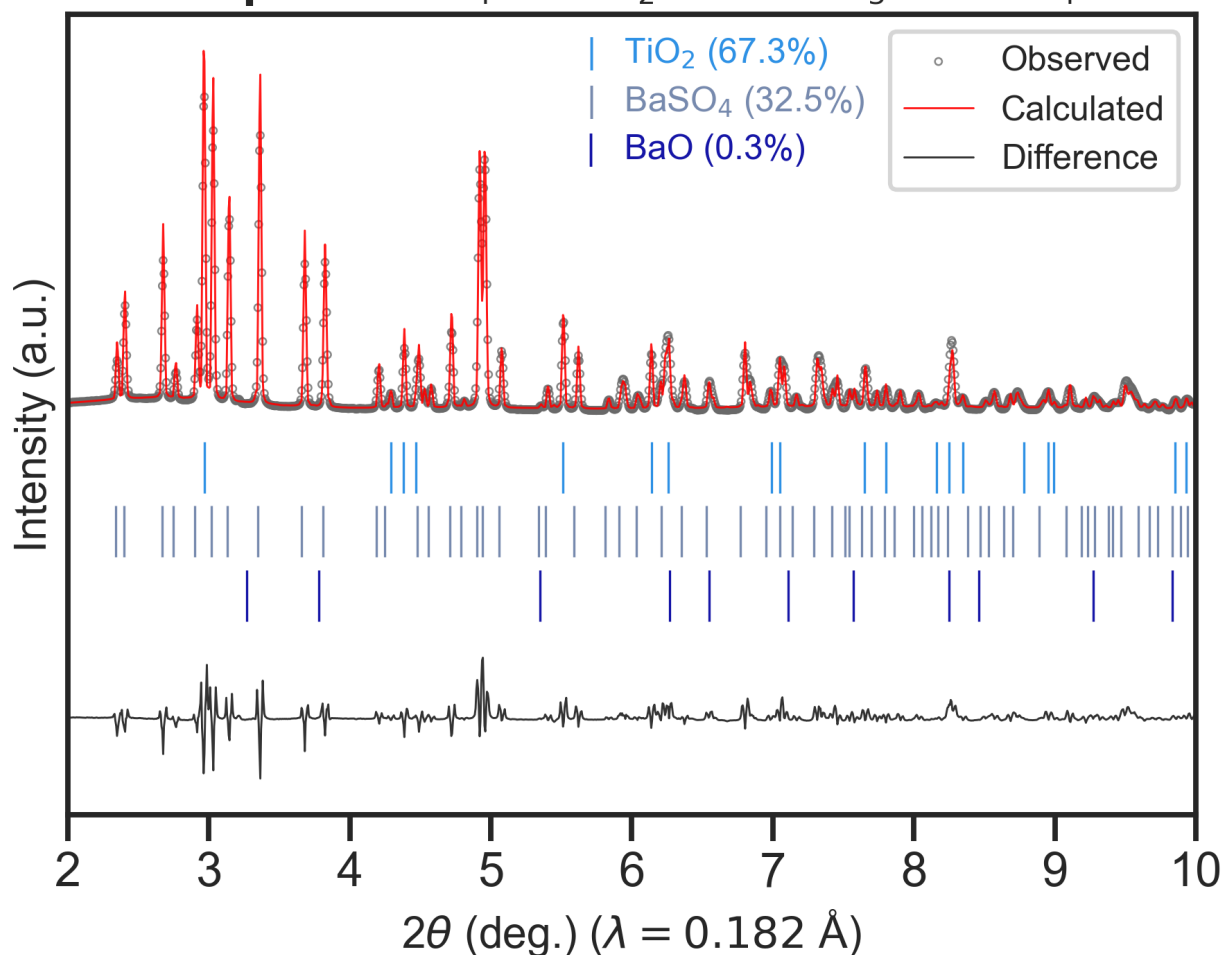
Figure S10: Selected Rietveld refinement for Experiment 7. The observed pattern represents *ex post facto* SPXRD data captured following reaction at $T = 727$ ℃, which corresponds to the temperature with the highest $BaTiO_3$ yield. Phase fractions are shown in units of mole percent.

Figure S11: Selected Rietveld refinement for Experiment 8. The observed pattern represents *ex post facto* SPXRD data captured following reaction at $T = 603$ ℃. This value corresponds to the median temperature studied since $BaTiO_3$ was not observed at any temperature. Phase fractions are shown in units of mole percent.
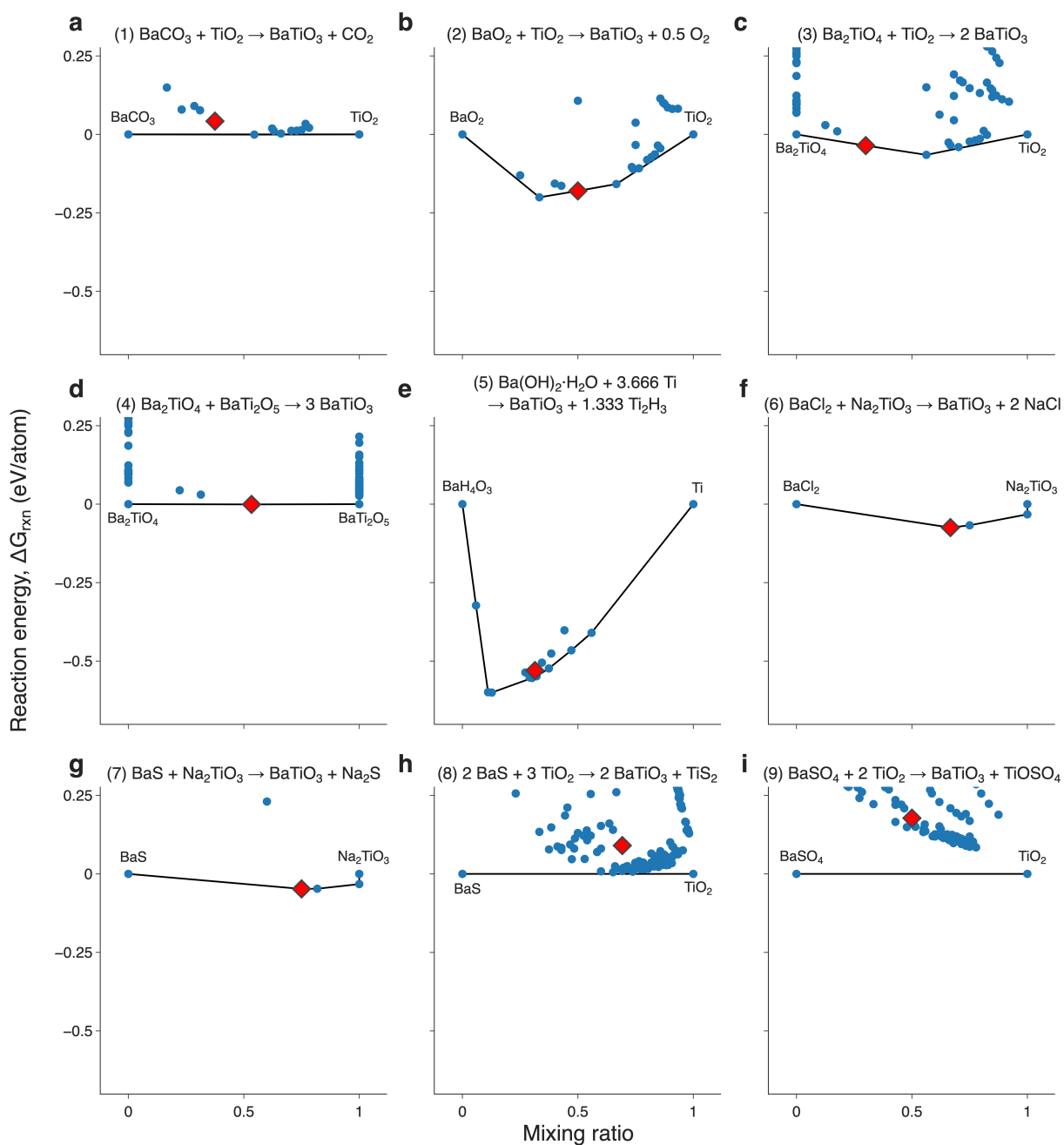
Figure S12: Selected Rietveld refinement for Experiment 9. The observed pattern represents *ex post facto* SPXRD data captured following reaction at $T = 594$ ℃. This value corresponds to the median temperature studied since $BaTiO_3$ was not observed at any temperature. Phase fractions are shown in units of mole percent.

# Interface reaction hulls for selected BaTiO$_3$ experiments

a (1) BaCO$_3$ + TiO$_2$ → BaTiO$_3$ + CO$_2$

b (2) BaO$_2$ + TiO$_2$ → BaTiO$_3$ + 0.5 O$_2$

c (3) Ba$_2$TiO$_4$ + TiO$_2$ → 2 BaTiO$_3$

d (4) Ba$_2$TiO$_4$ + BaTi$_2$O$_5$ → 3 BaTiO$_3$

e (5) Ba(OH)$_2$·H$_2$O + 3.666 Ti → BaTiO$_3$ + 1.333 Ti$_2$H$_3$

f (6) BaCl$_2$ + Na$_2$TiO$_3$ → BaTiO$_3$ + 2 NaCl

g (7) BaS + Na$_2$TiO$_3$ → BaTiO$_3$ + Na$_2$S

h (8) 2 BaS + 3 TiO$_2$ → 2 BaTiO$_3$ + TiS$_2$

i (9) BaSO$_4$ + 2 TiO$_2$ → BaTiO$_3$ + TiOSO$_4$

Reaction energy, ΔG$_{rxn}$ (eV/atom)

Mixing ratio

Figure S13: Interface reaction hulls for selected BaTiO$_3$ experiments. (a-i) Hulls for Experiments 1-9, as extracted from the full reaction network calculated during the synthesis planning workflow. Red diamonds mark the selected reactions of interest. All hulls are plotted on a uniform energy scale to facilitate comparison.

# All pairwise correlation plots between reaction metrics and experimental outcomes
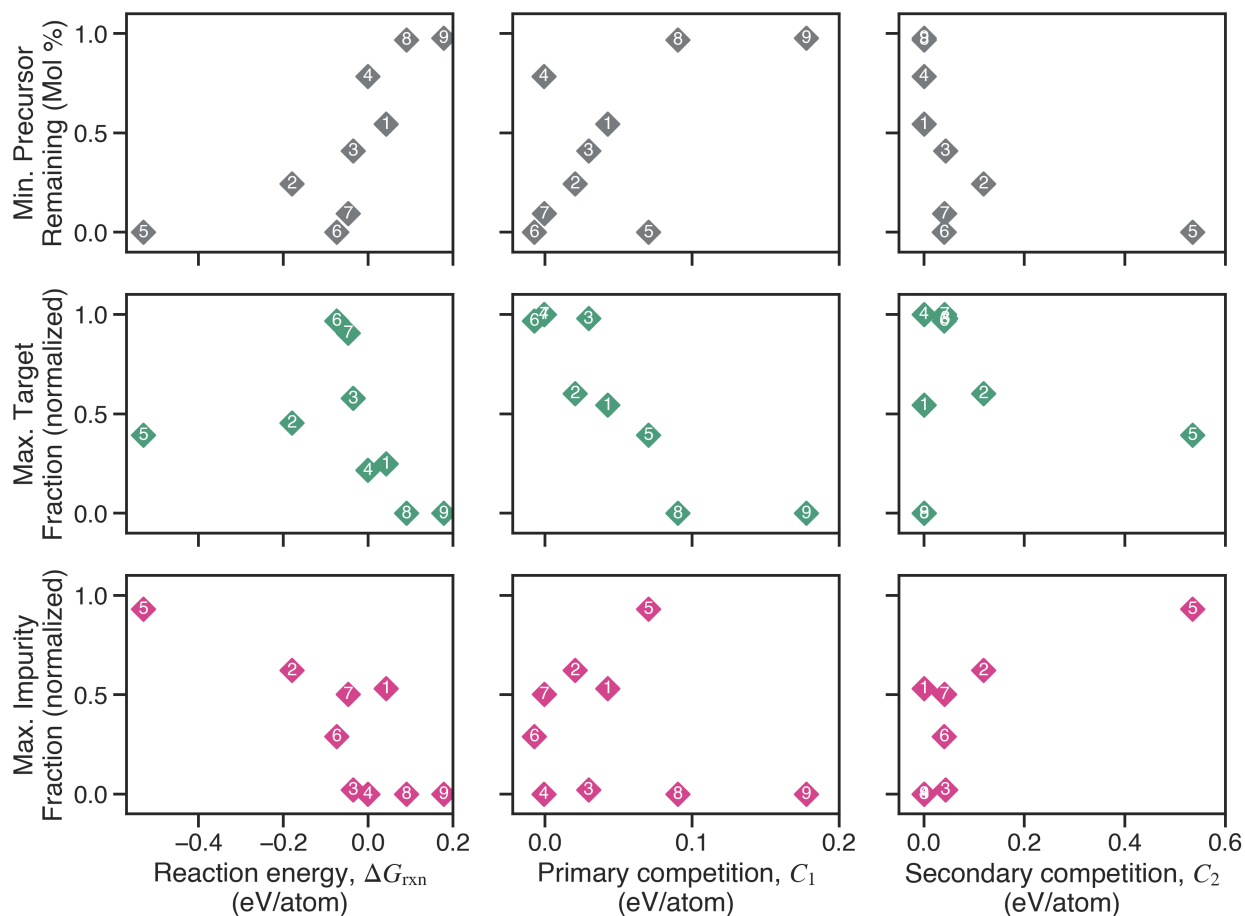
Figure S14: **Pairwise correlation plots between reaction metrics and experimental outcomes**. Plots of minimum remaining precursor (P), maximum target (T), and maximum impurity (I) formed as a function of reaction energy, primary competition, and secondary competition. The target and impurity plots have been normalized by the amount of precursor consumed ($1 - P$).

Correlations between selectivity metrics
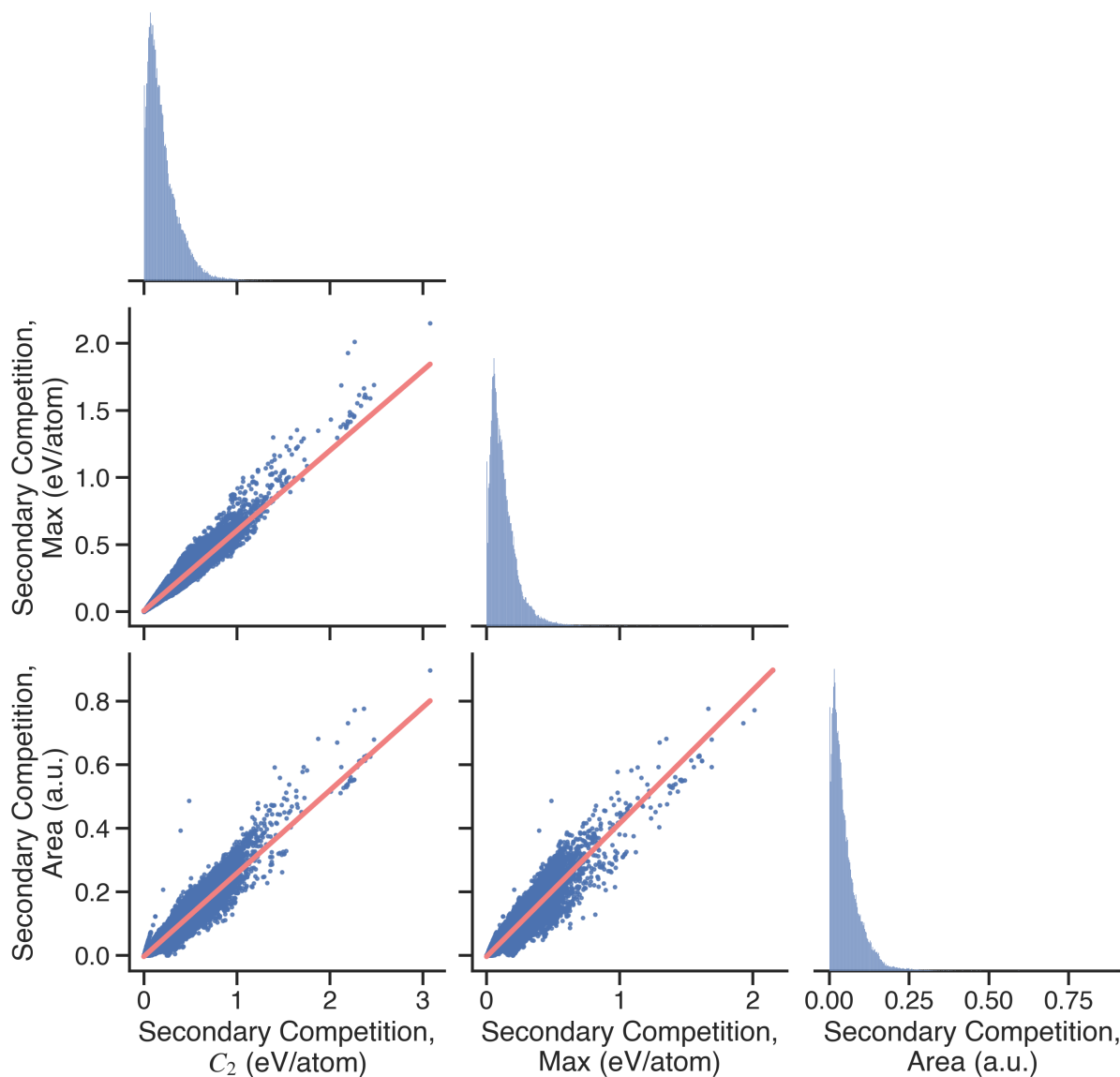
Figure S15: Pairwise correlations of alternative secondary competition metrics. Secondary competition (max) is defined as the sum of the energies of only the secondary reactions with the highest driving forces on either side of the target. Secondary competition (area) is the enclosed area of the interface reaction hull.
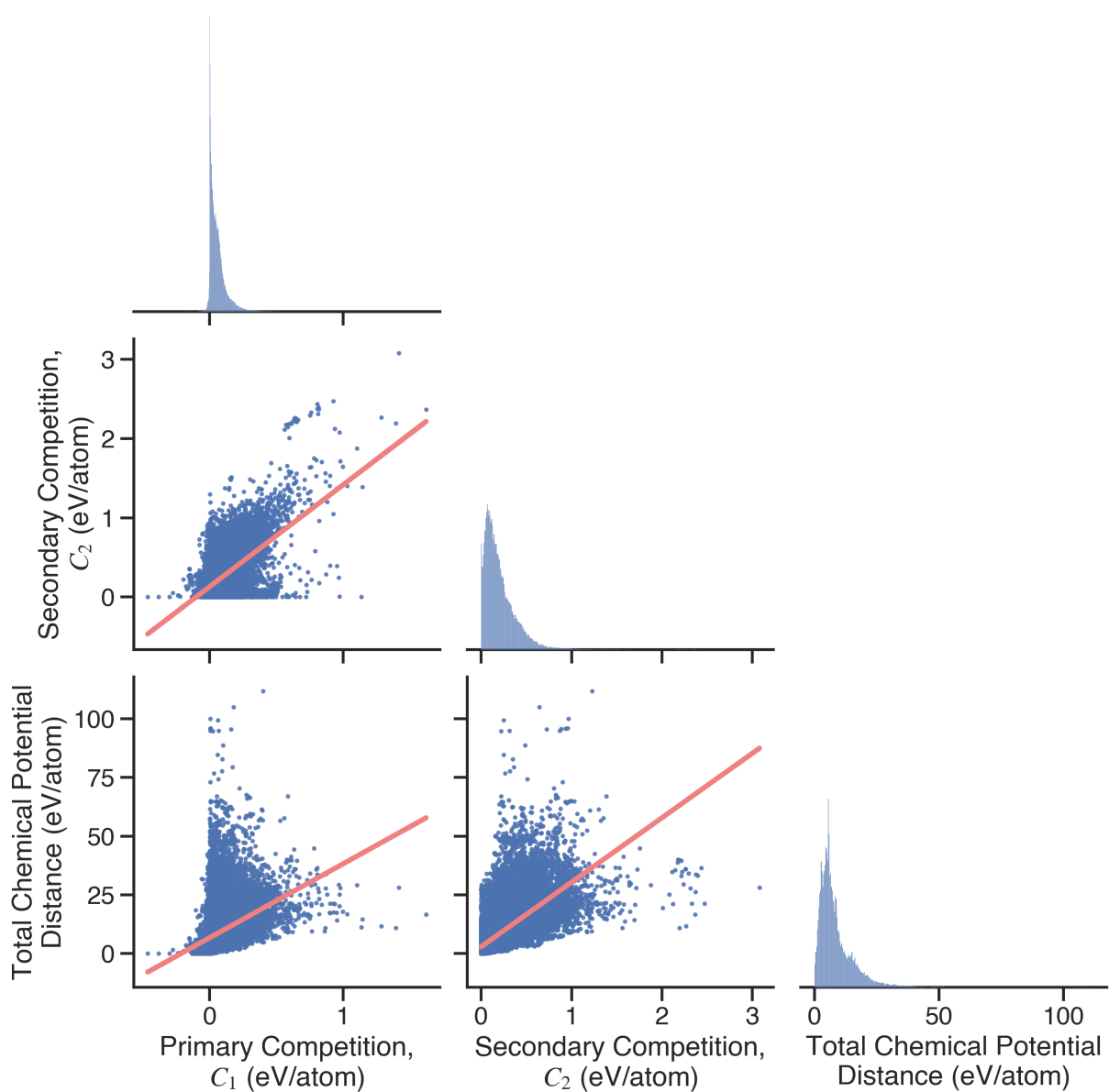
Figure S16: Pairwise correlations of selectivity metrics. The total chemical potential distance is calculated using the methodology outlined in Ref. S1.
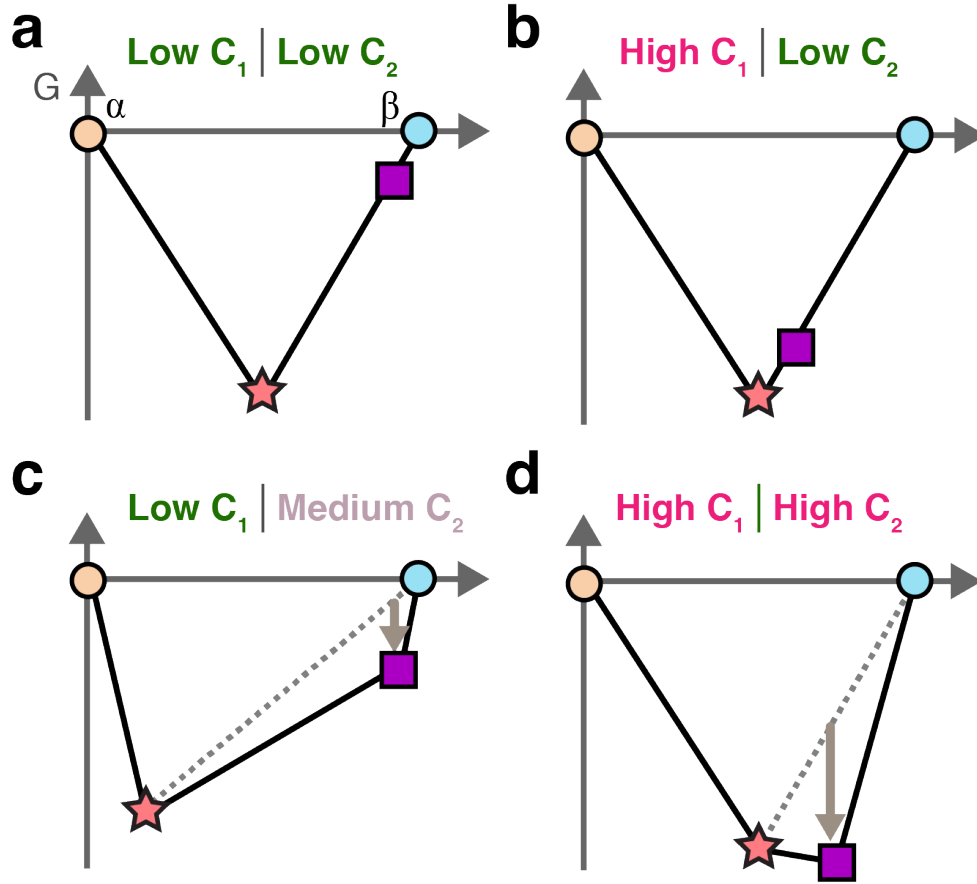
Figure S17: Examples of varying C_1 and C_2 indicating their partial correlation.
(a, b) When secondary competition (C_2) is low, primary competition (C_1) can either be low or high. c) However, when C_1 is low, C_2 can not be high due to the geometric constraints of the hull. d) There are no restrictions for both C_1 and C_2 to be high.

Formation energy correction for carbonates
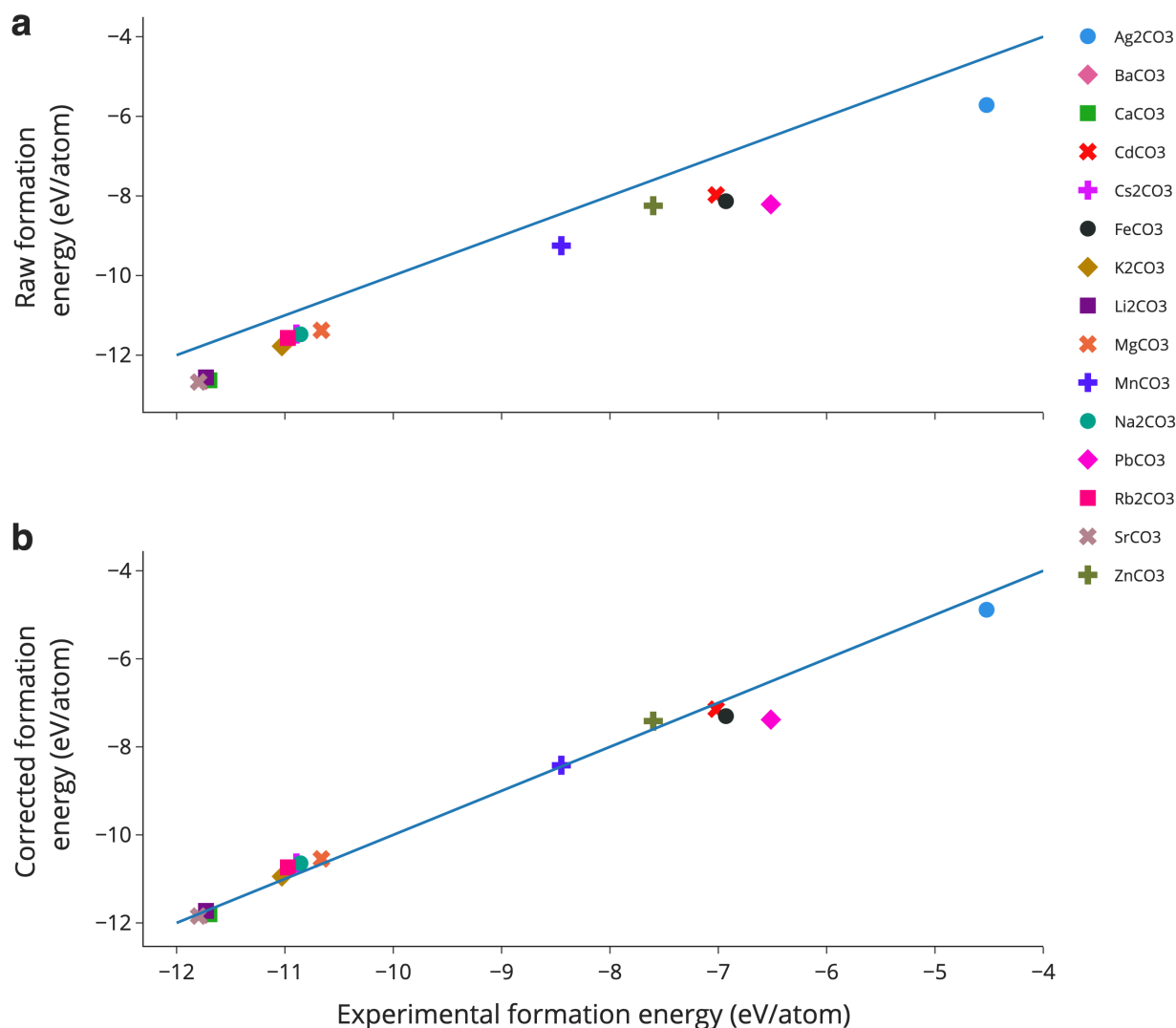
Figure S18: Fitting procedure for formation energy correction of carbonate compounds. (a) Carbonates exhibit a systematic negative shift in predicted Gibbs free energy of formation, ΔG_f ($T =$300 K) compared to the experimental values. (b) Energies of the same carbonate compounds after applying a fit energy correction of 0.830 eV/CO_3^{2-}.

# Gradient furnace setup, heating, and calibration
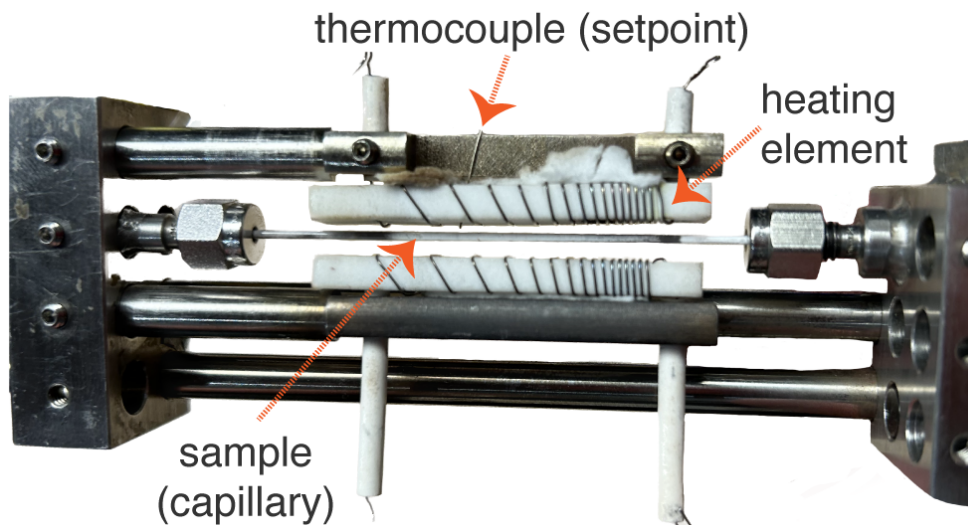
Figure S19: The gradient furnace used in all *ex post facto* SPXRD experiments.
The pictured device has the same specifications as described in Ref. S2. The heating wires
are wound with variable pitch, resulting in a wide temperature profile over the powder sample
(capillary tube). A thermocouple is used as a setpoint for determining the power supplied
to the heating elements.

Table S2: Reaction times for selected $BaTiO_3$ experiments. All times are shown in MM:SS
format (minutes/seconds). Temperatures correspond to gradient furnace setpoints of $T_H = 550$ ℃, $T_{L1} = 450$ ℃, and $T_{L2} = 400$ ℃.

Expt.	Precursors	Heat	Hold	Cool	Total	Temp.
1	$BaCO_3$-TiO_2	09:08	43:28	14:05	66:41	T_H
2	BaO_2-TiO_2	01:47	64:38	13:00	79:25	T_{L1}
3	Ba_2TiO_4-TiO_2	06:44	18:22	11:48	36:54	T_H
4	Ba_2TiO_4-$BaTi_2O_5$	12:07	14:30	12:35	39:12	T_H
5	$Ba(OH)_2 \cdot H_2O$-Ti	07:15	58:30	14:33	80:18	T_{L1}
6	$BaCl_2$-Na_2TiO_3	01:36	65:04	22:43	89:23	T_{L2}
7	BaS-Na_2TiO_3	02:14	61:26	28:16	91:56	T_{L1}
8	BaS-TiO_2	09:03	14:24	11:46	35:13	T_H
9	$BaSO_4$-TiO_2	07:15	25:54	13:46	46:55	T_{L1}
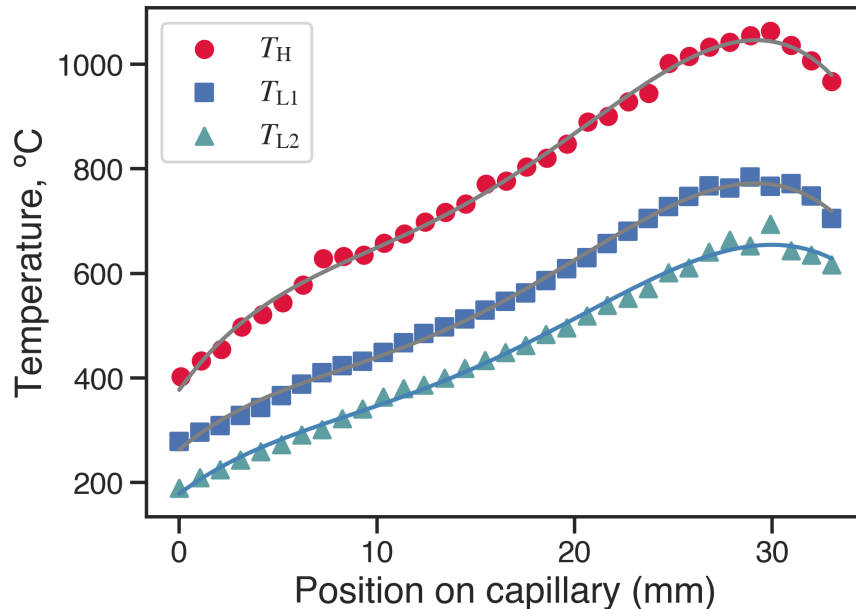
Figure S20: Measured temperature distributions along the length of sample capillary. Temperatures were determined by refinement of lattice parameters from NaCl/Si (T_{L1}, T_{L2}) or Al_2O_3/MgO standards (T_H). The calibration curves correspond to gradient furnace setpoints of $T_H = 550$ ℃, $T_{L1} = 450$ ℃, and $T_{L2} = 400$ ℃.

References

(S1) Todd, P. K.; McDermott, M. J.; Rom, C. L.; Corrao, A. A.; Denney, J. J.; Dwaraknath, S. S.; Khalifah, P. G.; Persson, K. A.; Neilson, J. R. Selectivity in Yttrium Manganese Oxide Synthesis via Local Chemical Potentials in Hyperdimensional Phase Space. *Journal of the American Chemical Society* **2021**, *143*, 15185–15194.

(S2) O'Nolan, D.; Huang, G.; Kamm, G. E.; Grenier, A.; Liu, C.-H.; Todd, P. K.; Wustrow, A.; Tran, G. T.; Montiel, D.; Neilson, J. R.; Billinge, S. J. L.; Chupas, P. J.; Thornton, K. S.; Chapman, K. W. A Thermal-Gradient Approach to Variable-Temperature Measurements Resolved in Space. *Journal of Applied Crystallography* **2020**, *53*, 662–670.