CHEMISTRY OF MATERIALS

pubs.acs.org/cm

Article

A Cellular Automaton Simulation for Predicting Phase Evolution in Solid-State Reactions

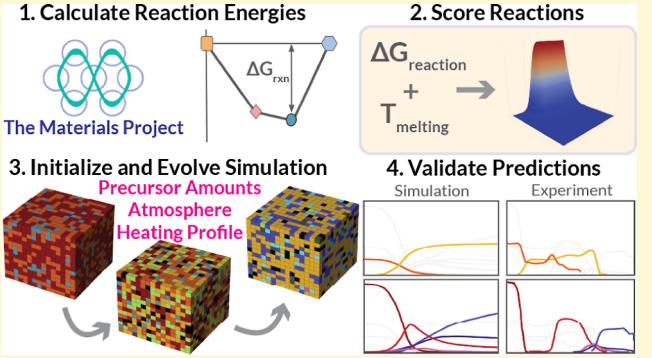
Max C. Gallant, Matthew J. McDermott, Bryant Li, and Kristin A. Persson*

Cite This: https://doi.org/10.1021/acs.chemmater.4c02301

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: New computational tools for solid-state synthesis recipe design are needed in order to accelerate the experimental realization of novel functional materials proposed by high-throughput materials discovery workflows. This work contributes a cellular automaton simulation framework for predicting the time-dependent evolution of intermediate and product phases during solid-state reactions as a function of precursor choice and amount, reaction atmosphere, and heating profile. The simulation captures the effects of reactant particle spatial distribution, particle melting, and reaction atmosphere. Reaction rates based on rudimentary kinetics are estimated using density functional theory data from the Materials Project and machine learning estimators for the melting point and the vibrational entropy component of the Gibbs free energy. The resulting simulation framework allows for the prediction of the likely outcome of a reaction recipe before any experiments are performed. We analyze five experimental solid-state recipes for $BaTiO_3$, $CaZrN_2$, and $YMnO_3$ found in the literature to illustrate the performance of the model in capturing reaction selectivity and reaction pathways as a function of temperature and precursor choice. This simulation framework offers an easier way to optimize existing recipes, aid in the identification of intermediates, and design effective recipes for yet unrealized inorganic solids *in silico*.



1. INTRODUCTION

The solid-state or "ceramic" method is a simple and ubiquitous technique for synthesizing inorganic crystalline solids in which powder precursors are heated to elevated reaction temperatures under controlled atmospheric conditions.[1] This method is used at both the small scale (e.g., research laboratories attempting the synthesis of new materials) and large scale (industrial manufacturing processes) to produce a wide variety of important functional materials such as battery cathodes (including $LiMnPO_4$[2] and $LiFePO_4$[3]), the ferroelectrics $BaTiO_3$,[4] $YMnO_3$,[5] and $BiFeO_3$,[6] and many superconductors, including $FeSe_{0.88}$,[7] $YBa_2Cu_3O_{6+x}$,[8] and MgB_2.[9] Despite the ubiquity of the method, no conventional system for designing or modeling solid-state synthesis recipes exists. Instead, recipes have long been designed primarily using expert knowledge (e.g., precursor selection from a common library or *via* phase diagram analysis) and heuristic guidelines.[10,11]

The difficulty in modeling solid-state synthesis reactions can be illustrated by drawing a comparison to organic molecular synthesis in which recipes can be generated by working backward from a desired product molecule to a set of known precursors *via* a series of mechanistically well-defined steps in a process known as retrosynthesis.[12] In contrast, high-temperature solid-state synthesis proceeds by spontaneous thermodynamic reactions, which lack clearly defined intermediates and reaction mechanisms. However, enabled by the recent rise of high-throughput density functional theory (DFT) calculations[13,14] and the databases generated by them,[15−18] several new automatable methods for designing solid-state synthesis recipes have emerged. These methods, which include measures for determining the synthesizability of a desired target,[19,20] metrics for comparing the selectivity of reaction recipes,[21,22] tools for extracting synthesis data directly from natural language,[23−25] and reaction networks that identify thermodynamically favorable pathways between precursor and target materials,[26] have yielded early success in guiding synthesis recipe design, despite being built on zero-temperature simulations of ordered crystalline structures. Furthermore, advances in autonomous synthesis have increased the throughput of synthesis experiments[27−29] and motivated the development of synthesis design algorithms that utilize experimental results to improve their planning.[30] While each of these methods provides an element of recipe design guidance, none of them allow for the quantitative prediction of

Received: August 15, 2024
Revised: November 18, 2024
Accepted: November 19, 2024

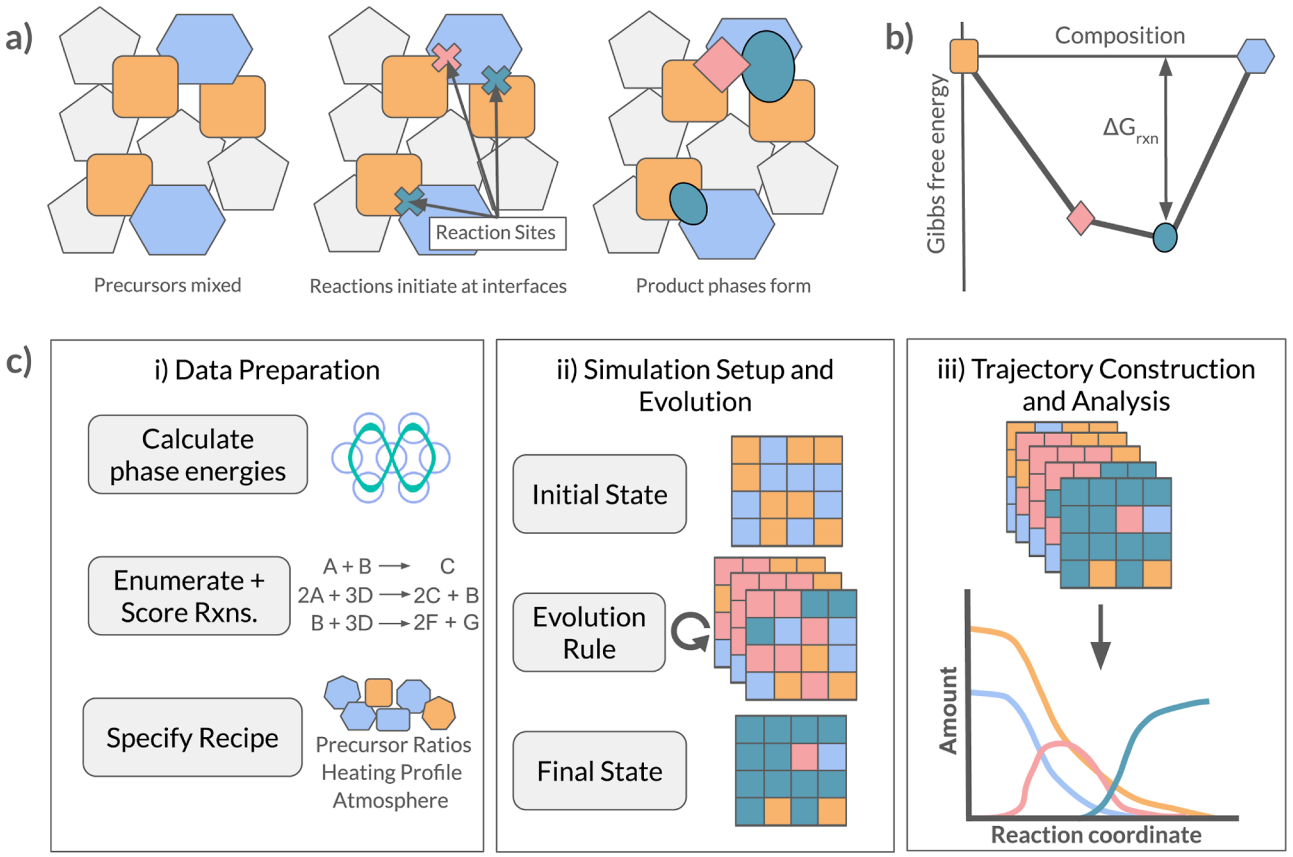


Figure 1. Modeling solid-state reactions with a CA. (a) Progression of the initial stages of a solid-state reaction occurring via the pairwise interface reaction model; (b) convex hull schematic representing the thermodynamics of reactions between two hypothetical solid precursor phases drawn as orange squares and light blue hexagons, with possible products given as the pink diamond and blue oval phases; (c) schematic illustrating the main stages of the simulation: (i) formation energies are obtained from the Materials Project and machine learning estimators are used to calculate the vibrational entropy part of the Gibbs energy of formation as a function of temperature and the melting point for each phase, reactions are enumerated and scored, and a recipe that defines the desired precursors, a heating profile, and a reaction atmosphere is specified; (ii) a random initial arrangement of particles is generated, and the evolution rule is repeatedly applied to simulate the reaction; and (iii) simulation steps are concatenated into a trajectory, which is analyzed to determine phase evolution over the reaction pathway.

the time- and temperature-resolved emergence and consumption of phases during the execution of a synthesis recipe.

Though no *a priori* simulation exists for predicting the progression of solid-state synthesis reactions, other reaction classes have been captured by simulation methods that are not neatly transferable to the solid-state case. For example, the kinetic Monte Carlo method is frequently used to model the evolution of species in gas or liquid phase molecular reactions (often in conjunction with reaction rates calculated using transition state theory[31,32]). This method assumes integer numbers of discrete particles that transform *via* reaction into other sets of discrete particles, and it assumes that these particles are available to interact with each other with no heed paid to their spatial arrangement.[33] These two assumptions do not hold for solid-state reactions; instead, solid phases transform in continuous amounts from reactant to product, and reactant particles do not move as freely as they do in liquid phase reactions (barring the presence of a molten flux or gas transport). Surface reactions have been successfully modeled by lattice Monte Carlo simulations to determine heterogeneous catalytic behavior, but they explicitly treat the motion of individual atoms.[34] This method is not feasible for modeling the evolution of the powder contents of a solid-state reaction vessel because the large number of atoms involved (often on the order of 10^{10} to 10^{20} atoms or more) leads to intractable computing requirements. Indeed, any atomistic method presents similar limitations. Finally, phase field models have been used to model ionic diffusion during solid-state metathesis reactions,[35] but these methods require significant assumptions about the form of the governing equations and explicitly known mobilities for each of the species involved. Such mobility values are not readily calculable, nor are they available in existing materials databases.

In light of these challenges, we present in this work a simulation framework (ReactCA) that predicts the time-dependent, quantitative evolution of phases over the course of a prescribed solid-state reaction as a function of the precursor ratio, heating profile, and reaction atmosphere. To achieve this, we leverage the cellular automaton (CA) formalism,[36] which offers a flexible framework for addressing the unique challenges posed by solid-state reactions. A CA is defined by a grid of sites, each of which is assigned a state value. At each step in the evolution of the automaton, the state in each site is updated according to its own current value and the states of the sites neighboring it (the "neighborhood"). The specific nature of the state values and the rule governing evolution (the "evolution rule") can be chosen to best suit the simulation problem at hand. As a result of this flexibility, cellular automata have been used in materials science and chemistry to model a variety of processes, including grain

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX

growth, crystallization, and surface adsorption/desorption.[37,38] Due to the fundamentally spatial nature of the neighborhood, the flexibility of the evolution rule, and the rapidly advancing theory of solid-state synthesis, the CA structure is a natural choice for modeling these reactions.

The simulation framework described here utilizes the zero-temperature thermodynamic properties of ordered crystalline compounds from the Materials Project as its primary input data. Importantly, some of the compounds we simulate here (and many of those found in the Materials Project) can accommodate disorder, which would increase the entropy of these phases and affect the energetics of the reactions that form them, especially at higher temperatures. To date, however, no existing database of computed material properties contains rigorous representations of configurationally disordered materials and their entropy. As a result, for the ordered materials in the Materials Project, we consider only the vibrational entropy contribution to the Gibbs formation energies as estimated by the machine learning method of Bartel et al.[39] These estimates, in conjunction with the CA formalism, are used to capture the thermodynamic and spatial features of solid-state reactions, in addition to some rudimentary kinetic effects based on machine learning estimates of phase melting points. Our framework offers new functionality in automated solid-state synthesis planning in that it enables the facile prediction of quantitative reaction outcomes *a priori* as a function of temperature profile, reaction atmosphere, and precursor choice. We envision our framework being used as a straightforward, easily implemented method 1) for testing hypothesized recipes before attempting them in the lab, 2) for implementing a digital twin in an autonomous laboratory designing its own synthesis recipes, and 3) for refining synthesis parameters when used in conjunction with optimization frameworks.

2. THEORY AND COMPUTATION

2.1. Cellular Automaton Model. The solid-state reaction CA simulation described herein (ReactCA) is constructed based on the pairwise model of solid-state reactions. This model states that solid-state powder reactions proceed predominantly *via* sequential reactions at pairwise interfaces (i.e., between only two solid species at a time). This model has its theoretical basis in the spatial geometry of the contact regions between particles and has been verified with *in situ* experiments.[8] The solid-state reaction process, illustrated in Figure 1a, proceeds *via* diffusion of atomic species driven by chemical potential differences across the interface between two reactant particles. As the reaction progresses, nuclei of one or more stable product phases form and grow at the interface, converting reactant material into product. Importantly, the local composition of the interface region is determined by the kinetic availability of reacting species and not constrained to reflect the overall composition of the precursor mixture. The first product phase to form is then a function of the "local" composition (as opposed to the overall composition) and the relative energetics of the possible product phases.[40]

The thermodynamics of the pairwise interface reaction model are conveniently represented by the convex hull construction, in which the Gibbs free energy is shown as a function of the mixing ratio (i.e., mixture composition) of the two precursor phases (Figure 1b). The interior points (i.e., pink diamond and blue oval) represent product phases that can form as a result of the reaction of the precursors. In two- or three-element systems, these points correspond to discrete compositions; however, for larger systems, the interior points can additionally correspond to balanced mixtures of two or more product phases. The vertical distance between the compositional axis and a product point is the change in the free energy of the corresponding reaction. The geometry of the reaction hull contains information about the behavior of a particular reacting pair, as illustrated by reaction selectivity metrics based on thermodynamics developed in ref 22. This pairwise reaction model and the interface hull underpin the simulation described here and, in particular, motivate the choice of the CA formalism, which naturally captures local interactions between neighboring entities.

The structure of the ReactCA simulation framework can be broken down into three stages (Figure 1c). The first entails the automated collection and calculation of relevant phase thermodynamics, an assessment of a score function for estimating relative reaction rates, and the specification of the reaction recipe. In the second phase, an initial state (or arrangement of phases on a grid) is produced, and then the evolution rule is repeatedly applied. Finally, the results of each application of the evolution rule are concatenated to form a trajectory that is analyzed to provide information about relative phase amounts at each time step.

2.2. Phase Data Acquisition. The input data for ReactCA are determined by the desired synthesis recipe that includes precursor ratios, a heating profile, and a reaction atmosphere (currently gaseous atmospheres consisting of only a single element are supported, e.g., N_2, O_2, or Ar). The heating profile is defined by the user as a list of heating stages that each have a temperature and duration (specified by number of simulation steps). Once this recipe is defined, ordered crystal structures in the chemical system spanned by the reaction atmosphere and precursor phases are identified, and their calculated formation energies are acquired from the Materials Project. Note that these formation energies are calculated *via* zero-temperature DFT, while solid-state reactions occur at elevated temperatures. However, exact calculations of finite temperature formation energies are not available from existing high-throughput databases. To bypass this data deficiency, a machine learning descriptor given by Bartel et al.[39] is used to estimate the vibrational entropy contribution to the Gibbs energy of formation for ordered solid phases at each of the temperatures specified in the reaction recipe. The Gibbs free energies of formation for common liquid/gaseous phases are acquired from experimental thermochemistry data (NIST-JANAF tables).[41] Finally, the melting points of all phases are estimated using the graph neural network model from Hong et al.[42]

2.3. Reaction Enumeration and Scoring. With the phases and energies acquired from the Materials Project, the reaction-network[26] Python package is used to identify all stoichiometrically possible reactions and calculate the changes in Gibbs free energy associated with them at each of the specified temperatures. While no general strategy exists for estimating the rate of solid-state reactions, predicting the evolution of phases during a reaction necessitates a model for relative reaction rates. To accomplish this, a score, S, is calculated for each reaction at each temperature using a heuristic function (eq 1), which returns the relative likelihood of each reaction occurring.

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX

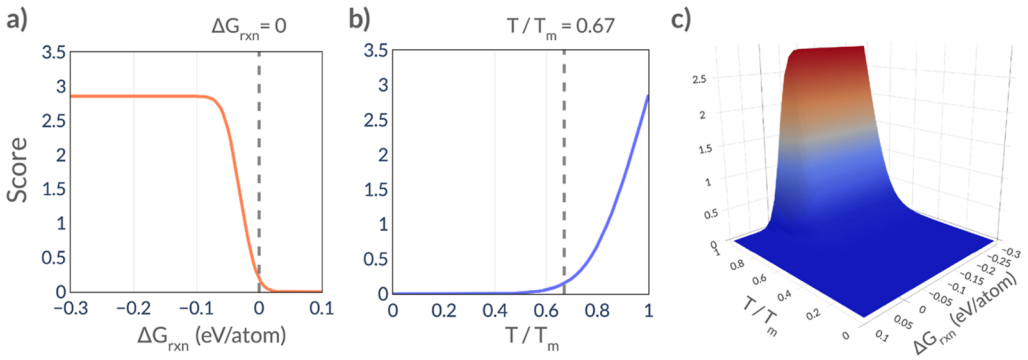


Figure 2. Scoring the likelihood of reactions as a function of reaction energy and temperature. (a) Score relationship with reaction energy; at a constant temperature above the Tamman temperature, endergonic reactions are vanishingly unlikely, while increasing exergonicity does not yield an infinitely increasing reaction rate, (b) relation with temperature (assuming a constant, negative reaction energy); reaction likelihood increases quickly above the Tamman temperature, (c) score relationship plotted as a surface function of both inputs.

$$S = \frac{1}{2}\ln\left[1 + \exp\left(a \times \left(\frac{T_{rxn}}{T_{m,reactant}} - b\right)\right)\right] \times [1 + \mathrm{erf}(-c(\Delta G_{rxn} + d))] \quad (1)$$

The score function described by eq 1 is composed of two primary terms: a softplus function term and an error function term. The shape of this function (as shown in Figure 2) captures: 1) the spontaneity of exergonic reactions, 2) the onset of reactions at temperatures equal to two-thirds of the melting point of the lowest melting point precursor (i.e., Tamman's rule), and 3) the increase of reaction rate with temperature. The scaling parameters $a = 14$ and $b = 0.8$ were chosen to shift/scale the softplus function such that its "onset" is around the Tamman temperature $\left(\frac{T_{rxn}}{T_m} = \frac{2}{3}\right)$ and the parameters $c = 35$ and $d = 0.03$ were chosen to shift/scale the error function such that it is centered on the region just below $\Delta G = 0$ eV/atom. Other values near the ones shown here for a, b, c, and d were experimented with, but the variations did not significantly alter the simulation outcomes.

The softplus function was chosen to encode Tamman's rule because of its "soft" activation (i.e., above the Tamman temperature). The error function was chosen to encode spontaneity because it behaves as a dial that abruptly "ramps up" for exergonic reactions. While these effects could also have been encoded using piecewise functions (e.g., a rectified linear unit in place of the softplus function or Heaviside function in place of the error function), we opt for smooth alternatives that "smear" the onset of each effect over a range of values. This smearing allows for a degree of accommodation for uncertainty in our input Gibbs energy and melting point estimates. Most importantly, the scoring function can easily be updated to accommodate more sophisticated functionality, e.g., based on kinetics and local availability of reactive species.

2.4. Simulation Evolution. After phase data are collected and reactions are enumerated and scored, an initial simulation state is defined. The simulation box for this automaton is a three-dimensional region of space subject to periodic boundary conditions and discretized into a grid of cubic cells. To establish an initial state, each cell is randomly assigned a phase occupancy according to the precursor ratios given by the reaction recipe, along with a volume equal to 1.0. We assign no scale or unit to this value because only the ratio of the volumes of neighboring cells is relevant to this simulation. This value is used in achieving conservation of mass and should not be interpreted as a literal measure of the physical extent of the simulation. After the initial state is established, the simulation evolves according to an evolution rule, which encodes reaction behavior. An animated visualization of the evolution process for a synthesis recipe for $YMnO_3$ (discussed later in this text) is provided in the Supporting Information.

The evolution rule determines the phase occupancy and volume value of the selected cell in the next simulation time step. It is applied to a single cell at a time, selected at random, meaning that this simulation is an asynchronous CA (as opposed to a standard CA, in which every cell is updated simultaneously).[43] This rule ensures that one of two actions occurs: 1) a swap or 2) a reaction. These actions are illustrated in Figure 3. The definitions for the simulation state and shape, along with the evolution rule, were implemented using the pylattica Python package.[44]

2.4.1. Action 1: Melted Phase Swap. If the current reaction temperature is above the melting point of the phase in the selected cell, then the state of the selected cell is swapped with one of its neighbors, chosen at random, as shown in Figure 3. To accommodate uncertainty in the estimation of the melting point, the onset of the likelihood of this swap is smeared over a range of relative temperature values. Specifically, the swap likelihood begins ramping up as a function of reaction temperature at $T_{rxn} = 0.8T_m$, increases to a 95% probability when $T_{rxn} = T_m$, and reaches a 99% probability at $T_{rxn} = 1.2T_m$. This behavior is shown graphically in Figure S1. The swapping motion facilitates movement of the reaction vessel contents and can capture heightened reactant movement during flux-mediated reactions, in which the presence of a liquid phase makes reactants more able to access each other. This is crucial for capturing more realistic reaction dynamics in many solid-state reactions.

2.4.2. Action 2: Reaction Progression. If the phase occupying the selected cell is determined to be a solid (i.e., it has a melting point higher than the current temperature), a reaction is selected. In this step, the reaction library is consulted to identify possible reactions between the selected cell and its neighbors. Reactions between the phase in the selected cell and the reaction atmosphere are also considered. From this list of possible interactions, a reaction is chosen randomly with a probability that is proportional to its score obtained from eq 1. As a result, reactions with higher scores occur more frequently than reactions with lower scores. This

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX

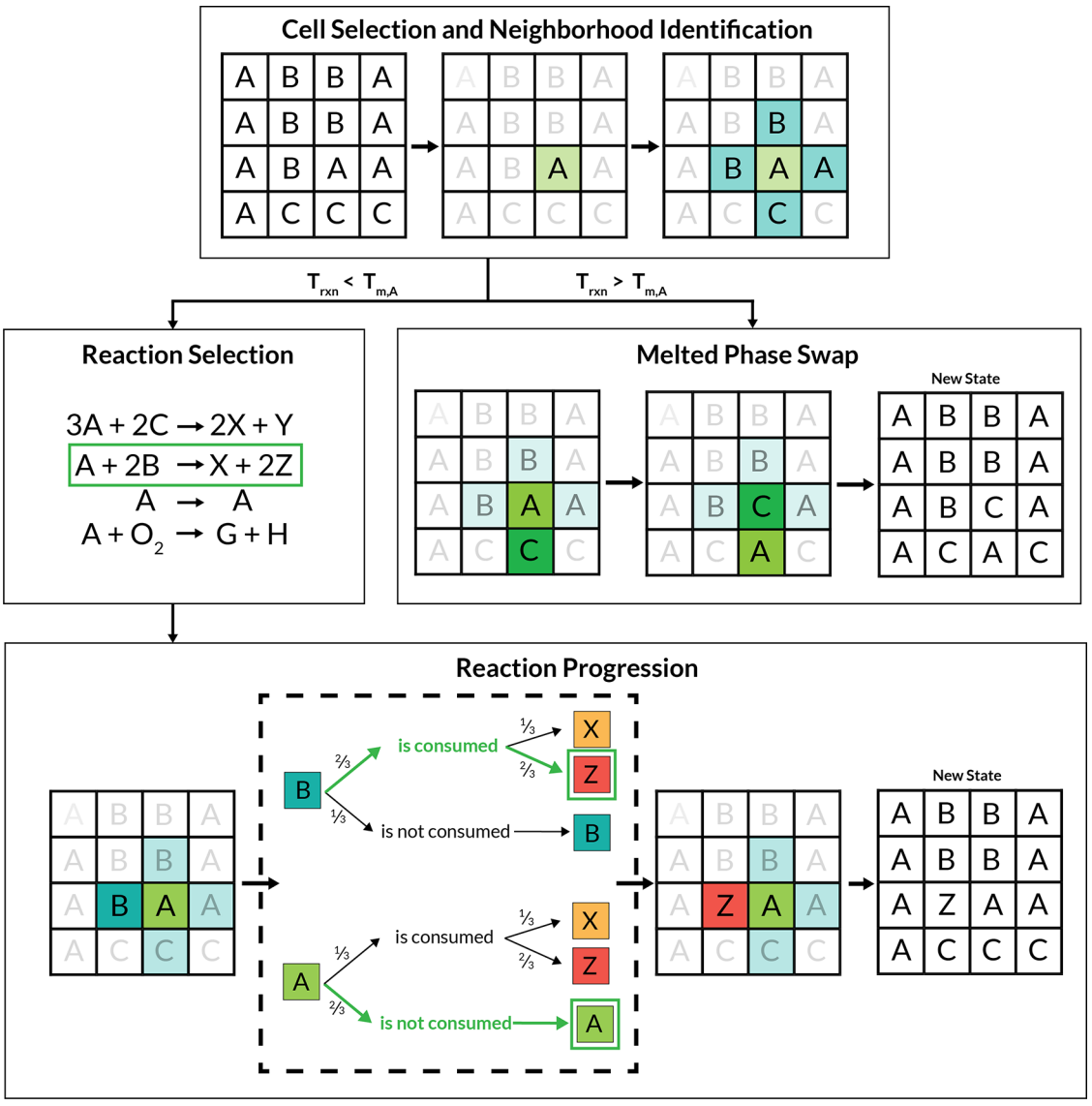


Figure 3. Evolution of phases in the simulation according to the evolution rule. In the top panel, a cell is randomly selected (pale green) from the simulation, and its neighbors are identified (teal). Next, if the simulation temperature is above the melting point of the phase in the selected cell, the Melted Phase Swap action occurs (middle-right). If not, reactions between the selected cell and its neighbors are enumerated and a reaction is randomly chosen using the reaction scores as probabilities (middle-left). Finally, each of the reacting cells' phases are replaced (or not) according to probabilities given by the stoichiometric coefficients of the selected reaction. These probabilities are indicated by the small fractions decorating each arrow in the bottom panel. Note that the CA implemented in this work uses a three-dimensional simulation state, but only two dimensions are shown here for clarity.

scheme has the net effect that higher-scoring reactions proceed faster. Once a reaction is chosen, the reaction proceeds at each of the reacting cells; note that only a single cell is involved if the other reactant is contained in the atmosphere (e.g., gaseous O_2). This procedure involves several steps (Figure 3):

1. A probability distribution over the reactants is constructed. The stoichiometric coefficients taken from the reaction are used as the weights in this distribution.
2. A random draw from the resulting distribution is performed. If the resulting phase matches the phase of the reacting cell, the process proceeds to the next step. If it does not match, the step ends, and the cell is left unchanged.
3. If the reaction proceeds, a second distribution is constructed over the reaction products (again using their stoichiometric coefficients as weights).
4. A draw from this distribution is used to select a product phase.
5. The reacting phase is replaced with the product phase.
6. The volume of the cell is scaled according to the ratio between the volume of the products and the reactants.

Importantly, the process described above utilizes probability distributions over the stoichiometric coefficients of the reaction to maintain conservation of mass within the automaton. For a given reaction, the coefficients of the reactants provide the probability that each will be consumed during the occurrence of that reaction. This ensures that the reactants are consumed at the correct rate relative to each other. The coefficients on the products of the reaction provide the probability that each one will be produced by a given occurrence of the reaction, similarly ensuring that the products of each reaction are produced at the correct rate relative to

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX

each other. Finally, scaling the volume of the simulation cell after its contents have been replaced during a reaction ensures that the correct amount of the product phase is produced relative to the amount of reactant consumed by the reaction. A more detailed explanation of this process in conjunction with an example is provided in the Supporting Information.

2.5. Trajectory Construction and Analysis. A simulation run typically entails hundreds of thousands of applications of the evolution rule described above. When the simulation is complete, the results are concatenated into a trajectory that can be analyzed to understand the reaction pathway as a series of steps and discrete intermediate species. Because ReactCA relies on random draws from probability distributions over the possible actions, reactions, and product phases, several simulations are run in parallel, each utilizing a different random starting state. Though the choice of starting state does not affect the qualitative outcome of the simulation, each trajectory is characterized by differing fluctuations and represents a unique sampling of the distributions in the automaton. As shown in Figures S2 and S3, when a sufficiently large simulation box is used, the standard deviation of the maximum and final mass fractions attained by each phase across a set of parallel trajectories is reduced to less than 1%. To construct the final result, the individual outputs of these parallel simulations are ensemble averaged to yield an overall trajectory. An illustration of the degree of fluctuation between individual trajectories is shown in Figure S7, where six separate and randomly initialized trajectories for a $YMnO_3$ synthesis recipe (described in detail in the Results and Discussion) are plotted. While the precise amount of each phase varies between the trajectories at each time step, the qualitative features of the prediction (in terms of major intermediates, their relative amounts, and the order in which they appear) do not significantly vary in this example (or in any example we have observed).

3. RESULTS AND DISCUSSION

To test the efficacy of ReactCA in describing real solid-state reactions, we apply it to several case studies selected from the literature where high-quality *in situ* phase evolution data are available. Each of the phase prevalence plots shown here was produced by averaging together six individual trajectories, which each evolved from a different randomly initialized starting state. For each case, we compare the predicted and observed final products of the synthesis reaction as well as the appearance (or disappearance) of intermediate and impurity phases.

3.1. Product Selectivity in $BaTiO_3$ Recipes. Barium titanate is a well-known multiferroic material with a significant body of synthesis literature. While there are a number of well-known recipes for producing this material, we refer to the recent solid-state reaction selectivity study of ref 22, which tested and compared nine different $BaTiO_3$ synthesis recipes characterized over a range of temperatures with synchrotron X-ray diffraction (XRD). A selection of these recipes and the corresponding reactions are simulated here using ReactCA to illustrate the way reaction selectivity is expressed in a phase evolution prediction.

Selectivity was assessed in the aforementioned study (ref 22) according to two metrics: primary competition (C_1), which quantifies the likelihood of impurities forming from the reaction of precursors, and secondary competition (C_2), which quantifies the likelihood of subsequent reactions consuming the desired products after they form. In the case of both of these metrics, a lower value corresponds to a more selective reaction, that is, one that is more likely to form only the desired product phase. To illustrate the way that selectivity presents itself in ReactCA simulations, three recipes were selected: the conventional recipe (Recipe I—$BaCO_3$ and TiO_2), a recipe with improved primary selectivity but worse secondary selectivity (Recipe II—Ba_2TiO_4 and TiO_2), and a metathesis reaction with excellent primary selectivity (Recipe III—$BaCl_2$ and Na_2TiO_3). The primary and secondary competition scores reported by ref 22 for each of these recipes are shown in Table 1.

Table 1. Selected Experimental $BaTiO_3$ Synthesis Reactions and Their Associated Primary Competition Scores, C_1, and Secondary Competition Scores, C_2, Calculated in Ref 22

recipe	precursor pair	C_1 (eV/at.)	C_2 (eV/at.)
1	$BaCO_3 + TiO_2$	0.043	0.000
2	$Ba_2TiO_4 + TiO_2$	0.030	0.043
3	$BaCl_2 + Na_2TiO_3$	−0.007	0.040

3.1.1. Recipe I. The simulation results for Recipe I are shown in Figure 4a, and the corresponding experimental outcome is shown in Figure 4b. This reaction received a relatively high primary competition score but a perfect (zero) secondary competition score, suggesting that there were competing phases that could form from the original precursors but that if the desired products were formed, they would be unlikely to be consumed by any secondary reactions. In the corresponding experiment from ref 22, the $BaTi_2O_5$ phase formed as an impurity in conjunction with the product, $BaTiO_3$, at around 1100 K. The result of the reaction automaton simulation for this recipe, shown in Figure 4a, predicts the formation of the target phase $BaTiO_3$ as well as the $BaTi_2O_5$ impurity phase at the same onset temperature (1100 K). However, it also predicts the appearance of two additional phases (Ba_2TiO_4 and $BaTi_4O_9$). The three impurities grow at a rate similar to that of the desired product, $BaTiO_3$, an effect also seen in the experiment. This result illustrates the way that secondary competition appears in a reaction: competing phases form during the reaction of the precursor materials, but once the desired product, $BaTiO_3$, is formed, it is not consumed by any subsequent reaction. In addition to the four phases predicted to appear, there were 35 other accessible phases that were not predicted to appear by the automaton. Finally, we highlight that a long hold time at the highest temperature was required in the simulation in order for the product and impurity phases to form in such an amount that illustrated the selectivity, whereas in the experiment every temperature was held for an equal time. We opted to shorten the lower temperature stages because no apparent reactivity occurred during the simulation.

3.1.2. Recipe II. The second reaction chosen here was shown[22] to improve the selectivity of the first at the cost of lowering the driving force of the reaction by choosing compositional members toward the interior of the convex hull as precursors. The ReactCA simulation (shown in Figure 4c) predicts a majority $BaTiO_3$ formation, and the formation of the impurity $BaTi_2O_5$. In the corresponding experiment (shown in Figure 4d), $BaTi_4O_9$ is observed as an impurity, but the simulated $BaTi_2O_5$ is notably absent. This is discussed

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX

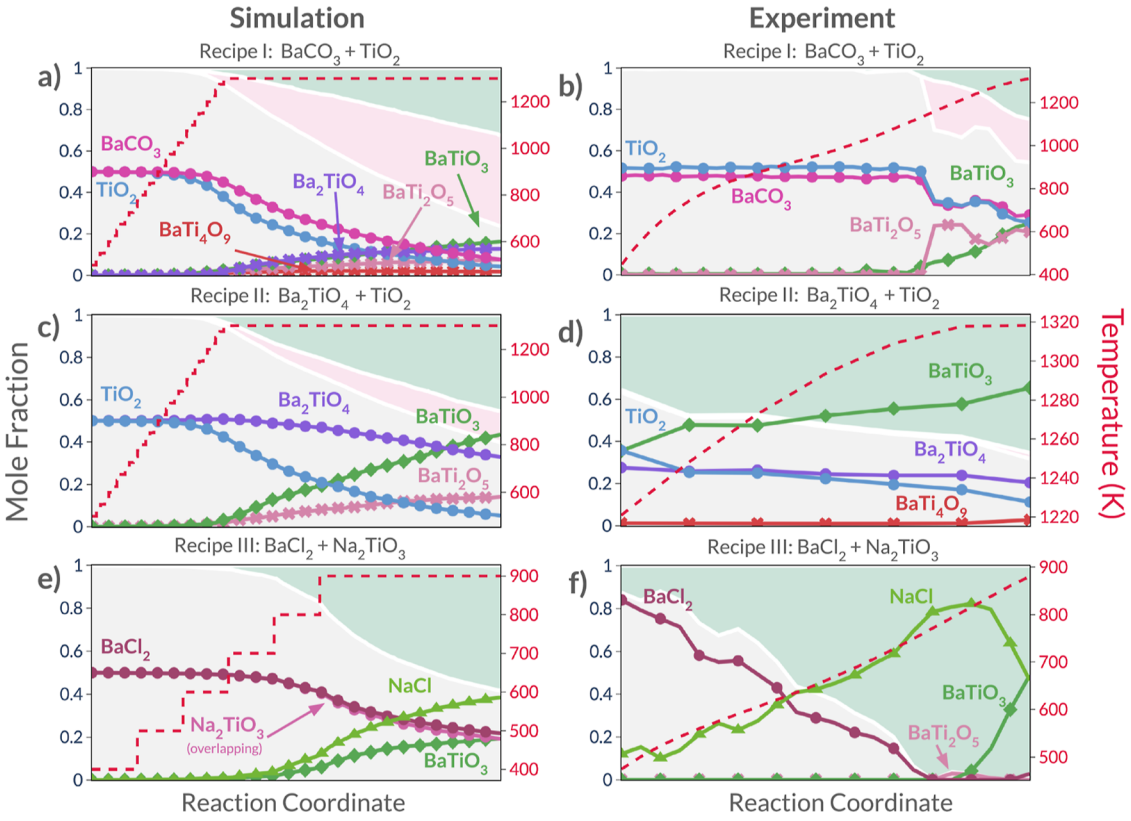


Figure 4. Simulated (left) and experimental (right) reaction evolution plots for the selected $BaTiO_3$ recipes. In each of these plots, the *x*-axis corresponds to the reaction coordinate, and the *y*-axis corresponds to mass fraction. The background of each plot is colored according to the amount of precursor (gray), impurity/intermediate (pink), and target/byproduct (light green) over the course of each simulation or experiment. Each of the traces represents the amount of each phase during the reaction. Traces marked with circles correspond to precursor phases, those marked with exes correspond to intermediate or impurity phases, and those marked with diamonds correspond to $BaTiO_3$, the expected target product phase. Dashed red lines show the heating profile used for each simulation or experiment. Experimental results were reproduced using data from ref 22.

below. In addition to the two phases predicted to appear, there were 30 other accessible phases that were not predicted to appear by the automaton. Additionally, we note that in the original experimental results for this recipe, no data were reported for phase amounts during the temperature ramp-up period. As a result, the experimental results shown here correspond to the high-temperature hold region in the simulation result. This also explains the fact that $BaTiO_3$ is already present at the left most point in the experimental phase evolution plot for this reaction.

3.1.3. Recipe III. The final $BaTiO_3$ reaction selected for simulation is a metathesis reaction using $NaTiO_3$ as the Ti source and $BaCl_2$ for the Ba source. In the original work, this reaction was selected for its strong exergonicity and its strong selectivity scores. The high selectivity of this reaction is on display in the ReactCA prediction, and the prediction (shown in Figure 4e) is in good agreement with the experimental results (Figure 4f)—the dominant products are the intended metathesis products: $BaTiO_3$ and NaCl. The main discrepancy between the prediction and the experiment is that the automaton predicts no other Ba−Ti−O phase formation, while the experiment indicates the appearance of $BaTi_2O_5$, though it is only a trace amount. Besides the two phases predicted to appear, there were 53 other accessible phases, none of which were predicted to appear by the automaton. We also note that no data for $NaTiO_3$ were present in the original experimental XRD refinement and phase prevalence plot for this recipe. As a result, the mole fraction of $BaCl_2$ shown in the experimental result is higher than the true value. In fact, the authors of the experimental work used a 1:1 $BaCl_2$:$NaTiO_3$ precursor ratio, the same values used in the simulation. Finally, in this example, similar to Recipe I, our simulation required a longer hold time at the highest temperature to allow the product phases to grow such that they were easily visualized. This could be corrected in the future with an improved score function that better captures the low temperature reactivity and sudden increases in reaction rate that occur with temperature for these reactions.

Across these three reactions, the selectivity differences between the recipes are apparent in the results from the CA simulations. The size of the green regions in Figure 4 shows the overall trend from low selectivity (in the case of Recipe I), to increased selectivity (in the case of Recipe II) and finally to perfect selectivity (in the case of Recipe III). We also note a tendency for ReactCA to predict the appearance of unobserved impurity phases (particularly in the case of Recipe I, as shown in Figure 4a). This is a result of the evolution rule that samples reactions based on their calculated rates. The effect is especially strong for the Ba−Ti−O chemical system, which contains many phases with similar energetics and which have similar melting points (the two features that are used in ReactCA to calculate reaction rates). Finally, in the reactions shown here, we also highlight a tendency of the automaton to overpredict the accumulation of Ba-rich Ba−Ti−O ternary

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX

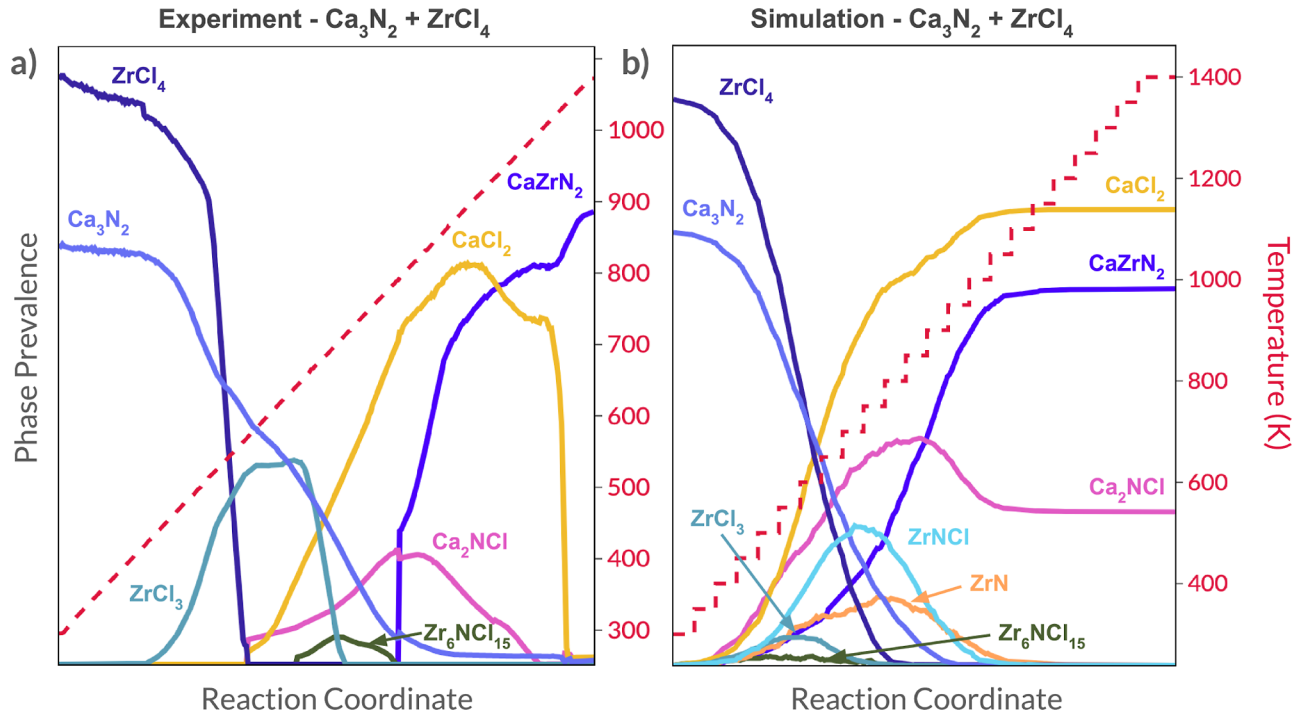


Figure 5. Resulting phase trajectories from simulation and experiment for $CaZrN_2$ synthesis. (a) Experimental phase trajectories for the reaction using Ca_3N_2 and $ZrCl_4$ as precursors and (b) simulated phase trajectories for the reaction using Ca_3N_2 and $ZrCl_4$ as precursors. In (a), the steep drop-offs of $CaCl_2$ and $ZrCl_4$ are caused by their melting and sublimating, respectively. Experimental results were reproduced using data from ref 46.

phases. For example, in the Recipe I simulation, the most prevalent byproduct is Ba_2TiO_4 (Ba-rich), but in the corresponding experiment, it is $BaTi_2O_5$ (Ba-poor). In the Recipe II simulation, $BaTi_2O_5$ (Ba-rich) is the primary impurity, but in the experiment, only $BaTi_4O_9$ (Ba-poor) appears. Finally, in Recipe III, the simulation predicts the formation of pure $BaTiO_3$ (Ba-rich), while in the experiment, small quantities of $BaTi_2O_5$ (Ba-poor) were also observed. In light of this observation, we hypothesize that - given the similar energetics of these compounds - the preferential formation of Ba-deficient compounds in experiments may be related to the kinetics of the ionic species across the reaction interface.[45] This discrepancy between simulation and experiment motivates future work to develop new reaction rate estimators using system-specific kinetic models, perhaps in conjunction with yet-unrealized high-throughput databases of kinetic calculations.

3.2. Intermediate Identification in $CaZrN_2$ Synthesis.

Ternary nitride systems provide a wealth of material discovery and synthesis opportunities. Recently, Rom et al. identified metathesis synthesis pathways that allowed them to produce the novel ternary nitrides $CaZrN_2$ and $CaHfN_2$.[46] Using *in situ* XRD analysis, they constructed trajectories for each phase present during their synthesis reaction, which used precursors Ca_3N_2 and $ZrCl_4$ to yield $CaZrN_2$. This trajectory is reproduced in Figure 5a. We performed a simulation for this reaction using a simulation box with a side length of 15 cells, a heating profile consisting of a ramp phase from room temperature to 1400 K, and a N_2 reaction atmosphere. The resulting phase trajectories are shown alongside the experimental results from ref 46 in Figure 5b. We note that in panel a) of this figure, the phase prevalence trace for Ca_4Cl_6O is excluded. This phase appeared in the experiment due possibly to either impure precursor material or reaction with the quartz ampule,[46] two effects that certainly represent practical synthesis considerations but are not relevant to the ideal environment represented by the automaton.

This simulation successfully captured the reaction pathway present in the experiment. The first intermediate phase to reach its peak in the simulation is $ZrCl_3$, which agrees with the early reduction of $ZrCl_4$ to $ZrCl_3$ in the experiment. The two intermediates that then follow in the experiment (a small amount of Zr_6NCl_{15} and significant Ca_2NCl) are also present in the simulation with the same relative prevalences. Finally, the simulation predicts the major product ($CaZrN_2$) and byproduct ($CaCl_2$) with confidence, though some unreacted Ca_2NCl also remains. This prediction was made from a set of 22 accessible phases, 15 of which were not predicted to appear by the automaton. In addition to this pathway, the simulation predicts the appearance of two phases that are not observed in the experiments, ZrNCl and ZrN. In the following analysis, we discuss the mechanisms by which these phases appear in the simulation and explain why those mechanisms may not have been active during the experiment.

ZrNCl appears in the simulation at roughly the same stage in the temperature trajectory as Ca_2NCl, and the three reactions that facilitated the bulk of its formation are shown in Table 2. The most frequent of these reactions, Reaction (1), may have incorrectly occurred in the simulation because ReactCA, in its current form, does not include sublimation. Reaction (1) consumes Ca_2NCl, which only forms in the experiment (Figure 5a) when the temperature has reached 600 K and $ZrCl_4$, the other reactant, has sublimated. As a result, the two reactants may never have been sufficiently available to one another for this reaction to occur in the experiment. In contrast, since ReactCA has no method for estimating

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX

Table 2. Most Frequently Occurring ZrNCl-Forming Reactions in the Automaton Simulation Shown in Figure 5b[a]

no	reaction	score (600 K)	$\Delta\Theta_{rxn}$ $\frac{eV}{atom}$	occurrences
1	$Ca_2NCl + ZrCl_4 \rightarrow 2CaCl_2 + ZrNCl$	0.165	−0.562	6683
2	$2Ca_3N_2 + ZrCl_4 \rightarrow ZrNCl + 3Ca_2NCl$	0.165	−0.636	5103
3	$Ca_3N_2 + 2ZrCl_4 \rightarrow 3CaCl_2 + 2ZrNCl$	0.165	−0.788	4533

[a]We use the notation $\Delta\Theta_{rxn}$ as opposed to ΔG_{rxn} to indicate that the relevant thermodynamic potential in this system is a grand potential with N_2 as the open species.

sublimation temperatures, $ZrCl_4$ remains available when Ca_2NCl appears, allowing Reaction (1) to proceed. This hypothesis is supported by another experiment by Rom et al., in which Ca_2NCl and $ZrCl_4$ were reacted directly as the initial solid precursors. In that experiment, both phases are present as solids, and ZrNCl appears as a prominent intermediate,[46] suggesting that Reaction (1) does occur if both precursors are present.

Reactions (2) and (3) in Table 2 consume the same precursors: Ca_3N_2 and $ZrCl_4$. In the experiment, however, Rom et al. propose that these phases instead react according to the following reaction

$$Ca_3N_2 + 6ZrCl_4 \rightarrow 6ZrCl_3 + N_2 + 3CaCl_2$$

$$S = 0.165,\ \Delta\Theta_{rxn} = -0.194\ \frac{eV}{atom} \qquad (T = 600\ K)$$

This reaction and the three reactions in Table 2 are all assigned the same score because they are highly exergonic (the thermodynamic component of the score function is maximized for all of them) and share the lowest melting point precursor, $ZrCl_4$ (so the melting point component of the score takes on the same value for all of them). Consequently, ReactCA does not differentiate the rates of these reactions, and they all occur with similar frequencies during the simulation. It is surprising that evidence for neither Reaction (2) nor Reaction (3) appears in the experiment, given their energetics (they are even more exergonic than the reaction proposed by Rom et al.), but there may be important differences in the kinetic accessibility of their product phases. In particular, the formation of the two binaries, $ZrCl_3$ and $CaCl_2$, (along with the release of gaseous N_2) from these two compositionally dissimilar precursors may be more kinetically facile than the formation of the nitrogen-containing ternary phases ZrNCl and Ca_2NCl. In support of this hypothesis, we note that Rom et al. observe the formation of a small amount of Zr_6NCl_{15} (Figure 5a), which could be interpreted as an incomplete incorporation of nitrogen while transforming of $ZrCl_3$ into ZrNCl. Additionally, while Ca_2NCl does appear in the experiment, Rom et al. suggest that its formation is facilitated by the reaction of more compositionally similar binaries, $CaCl_2$ and Ca_3N_2 (which is precisely how it is formed in the simulation). When ZrNCl does form, it occurs in the second experiment performed by Rom et al. (reacting Ca_2NCl with $ZrCl_4$). In this case, it may be that the presence of one ternary nitride (Ca_2NCl) leads to more facile formation of the other (ZrNCl), potentially by providing more favorable nucleation sites on account of the similar layered structures and shared $R\bar{3}m$ space group of the two phases. Considering these observations, the absence of ZrNCl in the experiment strongly motivates the development of improved kinetic estimations for reaction rates in ReactCA and suggests that such estimations might be based, in part, on compositional or structural features of the precursors.

The simulation also predicts the emergence and consumption of ZrN, an impurity that is not measured in either of the experiments. In their discussion, however, Rom et al. described the growth of the product, $CaZrN_2$, as facilitated by the slow growth of off-stoichiometric $Ca_xZr_{2-x}N_2$ starting from the ZrN rocksalt phase. In other words, the early product phase in their experiments *is* generated from ZrN, but the material at that point is likely heavily defective and stoichiometrically ambiguous, which may be the reason that no explicit ZrN phase appears in the XRD characterizations of the experiments. In contrast, no such defective or off-stoichiometric phase can be represented by ReactCA (which is limited to the ordered, crystalline, stoichiometrically exact phases currently available in the Materials Project). As a result, the explicit appearance and disappearance of crystalline ZrN are the best model the current version of the simulation can produce to represent the complex, continuous transformation in the experiment.

3.3. Recovery of Observed Reaction Pathways in $YMnO_3$ Synthesis. The multiferroic $YMnO_3$ has been the recent focus of a number of synthesis investigations into the effect of precursor selection, reaction atmosphere, and reaction temperature on both reaction pathways and the identity of the dominant product.[5,26,47] In the first of these studies, Todd et al. propose reaction pathways at work during the formation of $YMnO_3$ in a flux-assisted metathesis reaction and explain the lower reaction temperature required by their recipe in terms of the interplay between these pathways.[5] Building on this work, McDermott et al. were able to confirm using a reaction network that the suggested pathways were thermodynamically predicted by data within the Materials Project.[26] We use this example here to illustrate the ability of ReactCA to predict temperature-dependent reaction pathways and intermediate and product mass fractions and to provide insight into the interactions between the simultaneously occurring pathways.

The simulation for this reaction was configured to use a simulation box with a side length of 15 cells and a heating profile consisting of a ramp phase to 1300 K followed by a hold phase at 1300 K. We note that the peak temperature used here is slightly higher than the experimental maximum temperature of 1100 K. This choice was made in order to accommodate uncertainty in the melting point and vibrational entropy estimates. A view of the resulting trajectory for this simulation is shown in Figure 6a and a longer trajectory that includes the stabilization of the product phases is available in Figure S4. Additionally, an animation of the evolution of this simulation is provided in the Supporting Information. The overall result shown here predicts $YMnO_3$ as the dominant product phase and a number of intermediate phases, including YOCl, $Mn_8Cl_3O_{10}$, YMn_2O_5, and Mn_3O_4. In Figure 6b, a magnified view of the bottom of the trajectory is shown for those phases that never formed greater than 2.5% of the overall mass content in the simulation box. This prediction was made from a set of 105 accessible phases, 89 of which never formed during the simulation. Still, the multitude of phases present in this result illustrate the way that the reaction automaton samples many possible reaction pathways that can be traced from the initial precursor set.

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX

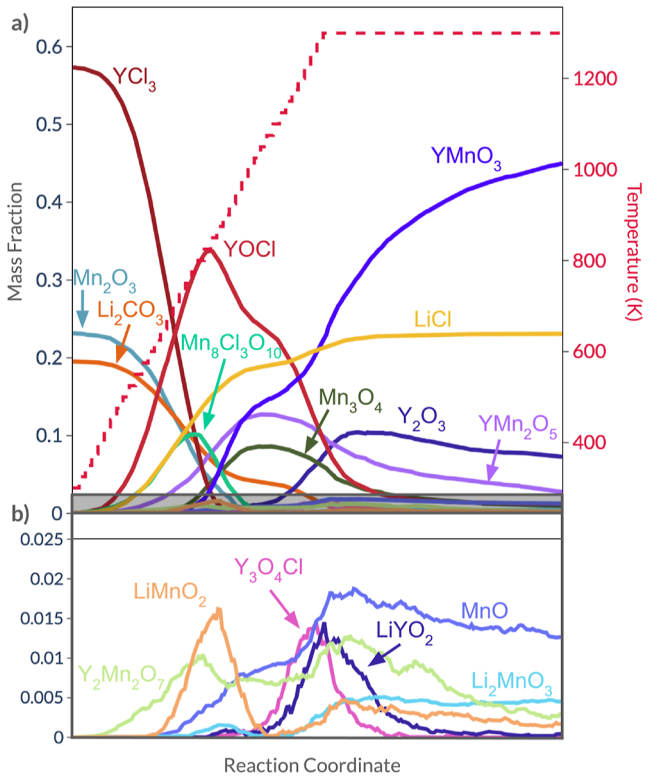


Figure 6. Resulting trajectory from simulation of the Li_2CO_3–YCl_3–Mn_2O_3 reaction. (a) Total trajectory and (b) a magnified view of all low-prevalence phases (≤2.5% by mass).

Amidst the complexity shown in Figure 6, the reaction pathways identified by Todd et al. are remarkably well predicted by the ReactCA simulation. From XRD refinements, Todd et al. calculated the trajectory of each intermediate phase and plotted them in groups according to cation. We reproduce these plots for the experimental data alongside similar plots generated from the simulation data in Figure 7. Every intermediate phase identified by Todd et al. is predicted by the reaction automaton. The relative amounts of these phases as well as the order of their appearance are also predicted with good accuracy, although Y_3O_4Cl appears in only trace quantities in the simulation. However, we note that previous work highlights the high degree to which both Y_3O_4Cl and YOCl accommodate defects and disorder.[48] Indeed, in another study, Todd et al. assert that the transformation between these phases proceeds through an off-stoichiometric $YO_{1+\epsilon}Cl_{1-\epsilon}$ phase.[48] Because our input data are limited to only the ordered, perfectly crystalline phases present in the Materials Project, we do not include the effects of defects and disorder. As a result, we neglect the likely significant contribution of configurational entropy to the stability of these phases. This omission may explain the underestimation of Y_3O_4Cl in this simulation result. Finally, we emphasize that the accuracy of these results is facilitated by the Melt-Swap action within the evolution rule. We show in Figure S5 that excluding this action from the evolution rule causes the reaction to stall after the appearance of the YOCl intermediate. This result aligns strongly with the experimental claim that the presence of a flux assists in transport during the reaction. We also note that the Tamman's rule heuristic portion of the score function is important for achieving this result. In Figure S6, the same recipe is simulated using a score function that excludes this heuristic (reducing it to a pure function of reaction thermodynamics). Using this alternative scorer, reactions occur speedily (and unrealistically) at room temperature, and Y_2O_3, a significant intermediate in the experimental pathway, appears only in trace amounts.

In addition to prediction of reaction intermediates, the ReactCA trajectory contains information about the specific reactions that yielded each phase. Of particular interest here is the formation mechanism for the product phase, $YMnO_3$.

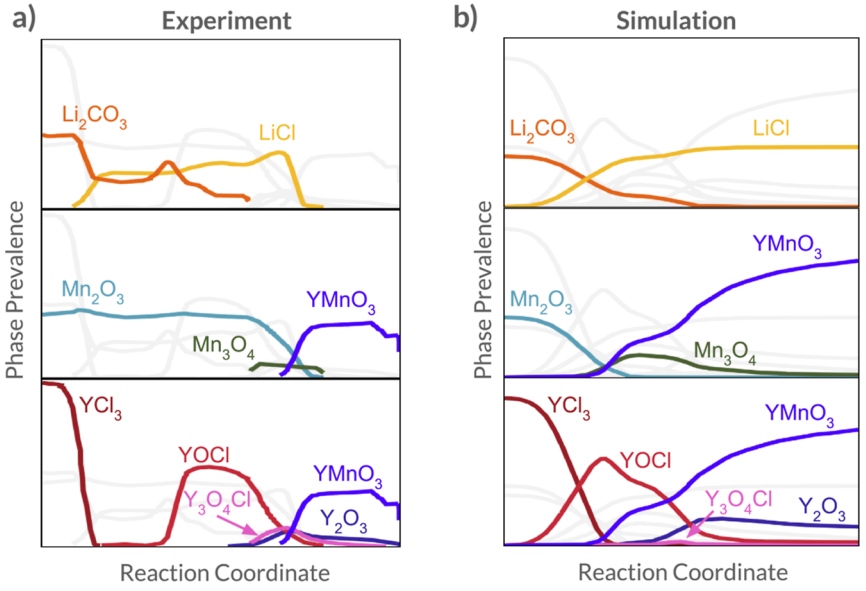


Figure 7. Reaction pathways extracted from (a) the experimental synthesis results from Todd et al.[5] and (b) the simulation trajectory of the reaction between Li_2CO_3, YCl_3, and Mn_2O_3. Top panels illustrate the early emergence and subsequent plateau of LiCl, the middle panels capture the emergence and then recession of the reduced Mn_3O_4 phase, and the bottom panel shows the ordering and relative prevalence of the three key Y–O–Cl intermediates. Note that LiCl disappears from the experimental phase trajectory in (a) (which was generated using XRD data) only because it melts. Experimental results were reproduced using data from ref 5.

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX–XXX

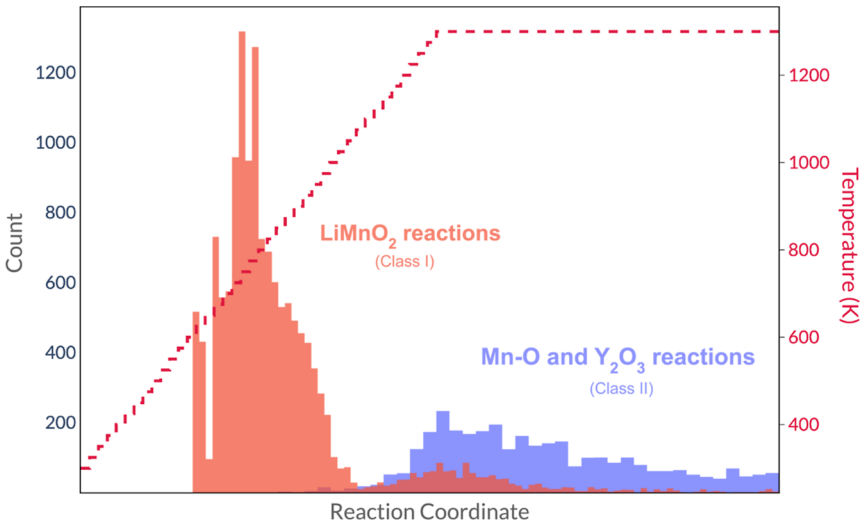


Figure 8. Histogram illustrating the dominance of different classes of $YMnO_3$-forming reactions over the course of the simulation. Counts shown in red correspond to reactions that involve the $LiMnO_2$ intermediate (the ternary metathesis route, i.e., Class I), and the counts in blue show reactions between the refractory Y_2O_3 and phases in the Mn−O chemical system (i.e., Class II).

Todd et al. propose dual mechanisms for the formation of this phase: first, the faster, lower temperature ternary metathesis reaction between $LiMnO_2$ and YOCl, and second, the higher temperature reaction between Mn_2O_3 and Y_2O_3 (the latter of which is formed in part by consumption of YOCl). Within the ReactCA trajectory, we identify two major classes of reactions that align with these two proposed reaction pathways. The first of these classes (Class I) is the reaction of $LiMnO_2$ with one of a number of yttrium-containing intermediates (the most frequently occurring of which is YOCl, followed by YCl_3). These reactions correspond to the low-temperature ternary metathesis step. The second of the two classes of reactions (Class II) producing $YMnO_3$ are reactions between the refractory Y_2O_3 and either Mn_2O_3 or Mn_3O_4. That reactions in the first class occur with any substantial frequency in this simulation is a surprising finding because the total amount of $LiMnO_2$ never exceeds 3% by mass, suggesting that the phase is consumed at a rate nearly equal to that at which it is produced. This may be a reasonable prediction, however, because the failure of this phase to accumulate in our simulation agrees strongly with the experiment performed by Todd et al., in which $LiMnO_2$ appears in only trace quantities (its XRD pattern is poorly resolved from Mn_3O_4, which implies that the data in Figure 7a suggests the appearance of only small quantities of both Mn_3O_4 and $LiMnO_2$ in the experiment).

By counting the frequency of each of these classes of reactions as a function of the reaction coordinate, we illustrate in Figure 8 that the first ternary metathesis reaction class dominates early in the simulation at lower temperatures and that the second reaction class, which consumes the refractory Y_2O_3, occurs later in the simulation at higher temperatures. This result reflects the mechanism that Todd et al. proposed based on their experimental observations: ternary metathesis dominates at lower temperatures, Y_2O_3 reactions proceed at higher temperatures, and indeed both mechanisms contribute to the formation of $YMnO_3$.

We emphasize that the analysis presented here represents an important increase in capability over the previous reaction network approach.[26] In particular, the reaction network requires a set of expected product phases as input, while ReactCA predicts the product without any prior target input. Additionally, the reaction network identified isolated pathways of a finite length at a single temperature at a time. In contrast, ReactCA can explore pathways of unlimited length (and allow them to interact) over a range of temperatures in a single simulation. Compared with the reaction network, these improvements both simplify the analytic process and broaden the range of reaction behavior that can be predicted.

In addition to the reaction intermediates identified by Todd et al., the ReactCA simulation predicts a number of other unobserved intermediates, most of which appear in the simulation at only trace levels (less than 1−2% by mass, as shown in Figure 6b). These unobserved intermediates reveal the alternate pathways that are present in the trajectory as a result of the automaton sampling many available reactions during its evolution. While we demonstrated earlier that investigating even low-prevalence intermediates, such as $LiMnO_2$, can yield insights into overall reaction pathways, the prevalence of a given intermediate in a ReactCA simulation generally reflects the relative amount predicted to appear during the synthesis reaction. In other words, while we do show that $LiMnO_2$ plays a role in the reaction that aligns well with the experimental hypothesis, its low prevalence suggests that it may not accumulate in significant quantities during the reaction. Similarly, because these other low-prevalence phases appear in only trace quantities, the ReactCA simulation should be interpreted as assigning a low likelihood to the appearance of those phases in experimental characterization.

Besides these low-prevalence unobserved impurities, two others (YMn_2O_5 and $Mn_8Cl_3O_{10}$) achieve significant amounts comparable to the predictions for the observed intermediates (>10% by mass). The first of these phases, YMn_2O_5, does not appear in the experiment originally presented by Todd et al., but is recognized in a later work by the same authors as a common impurity in the synthesis of $YMnO_3$ by this metathesis method.[47] The second of these phases, $Mn_8Cl_3O_{10}$, was recently synthesized by the solid-state method from precursors $MnCl_2$ and MnO_2[49] at 600 °C, hence its appearance is not implausible in this reaction. However, by examining the reactions that consume $Mn_8Cl_3O_{10}$ during the ReactCA simulation we find that its primary role is as an intermediate between reactants YCl_3 and Mn_2O_3 (which react

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX

to form it) and experimentally verified downstream intermediates (most significantly, $LiMnO_2$, Mn_3O_4, and LiCl). As a result, we hypothesize that $Mn_8Cl_3O_{10}$ functions similarly to ZrN in our discussion of the $CaZrN_2$ synthesis discussed above. That is, $Mn_8Cl_3O_{10}$ may be the best representation the simulation can provide of a process that is actually facilitated by highly defective or amorphous intermediates with a similar Mn−Cl−O composition that are not present among the ordered, crystalline, and stoichiometrically exact phases on the Materials Project. This uncertainty further emphasizes the necessity of future work to improve the ability of the automaton to navigate more sophisticated intermediate landscapes.

4. CONCLUSIONS

We present ReactCA, a new simulation framework based on the CA formalism for predicting the evolution of crystalline phases during the course of a solid-state reaction. This simulation utilizes thermodynamic data from the Materials Project and machine learning estimated melting points in conjunction with a cost function to assign reaction rates as a function of temperature. The evolution of the reacting material is determined by a rule based on the pairwise interface reaction model, which considers reactions between neighboring particles. The flexibility of the form of both the cost function and the evolution rule lends great extensibility, meaning that as new data become available from as of yet unrealized high-throughput methods or machine learning frameworks, both empirical rules and heuristics based on those data can be incorporated into ReactCA to improve its performance. We illustrate the current performance of the simulation framework using three case study systems, the first of which serves as a platform for viewing the relative selectivity of various reaction recipes and the importance of precursor choice with regard to product purity, and the second and third of which demonstrate the power of ReactCA in predicting likely reaction pathways, the temperature dependence of those pathways, and the order and amounts of intermediate phases that appear during the course of complex ternary metathesis reactions.

While we believe that ReactCA is of immediate utility in both "testing" reaction recipes before utilizing physical or monetary resources to experimentally execute them, we also foresee the simulation being used to determine recipe parameters using an optimization framework or to facilitate the autonomous design of synthesis recipes by acting as a digital twin for experimental synthesis in an automated lab. Still, there are many physical phenomena that are not considered by ReactCA. In particular, future work that will most significantly improve this simulation will be related to the intelligent incorporation of kinetic effects into the evolution rule and source data. The case studies presented in this work show that the current model likely overpredicts the number of ordered intermediates or impurity phases as compared with real synthesis experiments. Furthermore, including complex, noncrystalline intermediates such as off-stoichiometric and amorphous phases as well as additional phase transition mechanisms (e.g., sublimation or peritectic decomposition) provide future challenges. We foresee ReactCA growing alongside increasingly rich data generation capabilities in computational materials science to capture improved features of solid-state synthesis as the field matures, yielding faster and more accurate methods for predicting solid-state synthesis behavior.

5. METHODS

5.1. Thermochemistry Data. To prepare data for the simulation of the $BaTiO_3$, $CaZrN_2$, and $YMnO_3$ recipes, entries from the Materials Project from the Ba−Ti−O−C, Ba−Ti−O−Na−S, Ba−Ti−O−Na−Cl, Ca−Zr−N−Cl, and Y−Mn−Cl−Li−C−O chemical systems were collected. Additionally, a DFT structure relaxation calculation was performed at the GGA level of theory using parameters from the MPRelaxSet in atomate2[50] to obtain a formation enthalpy for Y_3O_4Cl, a phase known to appear during the synthesis of $YMnO_3$, but not present within the Materials Project. For each entry in these chemical systems, the machine learning descriptor by Bartel et al. was used to estimate the vibrational entropy contribution to the Gibbs energy of formation at 300 K.[39] Phases with formation energies greater than 30 meV/atom above the hull were removed at this temperature. This value for a metastability cutoff filter was chosen based in part on the work of Sun et al., which presents statistics on the metastability of compounds in the Materials Project and connects them to synthesizability.[20] We also removed compounds that were marked as "theoretical" on the Materials Project unless they appeared explicitly in the experimental results. In addition, several phases that have previously only been successfully produced using synthesis methods other than the solid-state method considered here were excluded. These are ZrCl (synthesized using high-pressure methods[51]), $ZrCl_2$ (synthesized via hydrogen disproportionation reactions[52]), CaN_2 (synthesized using high-pressure methods[53]), CaN_6 (synthesized in solution[54]), Ca_2N (synthesized via metallothermic reduction[55]), and Zr_3N_4 (synthesized via high-pressure methods[56,57] or ammonolysis[58]).

Data from the Materials Project was collected using the mp-api package and the possible reactions in each of these systems were enumerated using the reaction-network Python package.[26] The CA model was implemented using the pylattica Python package.[44]

■ ASSOCIATED CONTENT

Data Availability Statement

The implementation for this automaton is available at https://github.com/mcgalcode/rxn-ca. As implemented in the automaton repository, reaction enumeration was performed using the reaction-network package available at https://github.com/materialsproject/reaction-network.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.chemmater.4c02301.

Animated visualization of the $YMnO_3$ synthesis reaction simulation (MP4)

Swapping probability as a function of temperature; description of the reaction progression action; outcome convergence plots; extended trajectory for $YMnO_3$ reaction; $YMnO_3$ trajectory without melt-swap action; $YMnO_3$ trajectory without Tamman's rule score component; and individual trajectories for the $YMnO_3$ simulation (PDF)

Enumerated and scored reactions for the Ba−Ti−C−O chemical system (JSON), enumerated and scored reactions for the Ba−Ti−Na−S chemical system (JSON), enumerated and scored reactions for the Y−C−Mn−O−Li chemical system (JSON), enumerated and scored reactions for the Ca−Zr−N−Cl chemical system (JSON), serialized simulation recipe for $BaCO_3$−TiO_2 reaction JSON, serialized simulation recipe for Ba_2TiO_4−TiO_2 reaction JSON, serialized simulation recipe for $BaCl_2$−Na_2TiO_3 reaction JSON, serialized simulation recipe for $YMnO_3$ synthesis reaction JSON, serialized simulation recipe for $CaZrN_2$ synthesis reaction JSON, serialized simulation result for $BaCO_3$−TiO_2 reaction JSON, serialized simulation result for Ba_2TiO_4−TiO_2 reaction JSON, serialized simulation result for $BaCl_2$−Na_2TiO_3 reaction JSON, serialized simulation result for $YMnO_3$ synthesis reaction JSON, and serialized simulation result for $CaZrN_2$ synthesis reaction JSON (ZIP)

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX

AUTHOR INFORMATION

Corresponding Author

Kristin A. Persson − *Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Materials Science and Engineering, University of California, Berkeley, California 94720, United States*; Email: kapersson@lbl.gov

Authors

Max C. Gallant − *Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Materials Science and Engineering, University of California, Berkeley, California 94720, United States*; orcid.org/0009-0008-4099-6144

Matthew J. McDermott − *Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Materials Science and Engineering, University of California, Berkeley, California 94720, United States*; orcid.org/0000-0002-4071-3000

Bryant Li − *Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Materials Science and Engineering, University of California, Berkeley, California 94720, United States*; orcid.org/0000-0001-5482-509X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.chemmater.4c02301

Author Contributions

M.C.G.: conceptualization (lead); software (lead); and writing—original draft (lead). **M.J.M.**: conceptualization (supporting) and writing—review and editing (equal). **B.L.**: conceptualization (supporting) and writing—review and editing (supporting). **K.A.P.**: supervision; funding acquisition; and writing—review and editing (equal).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was primarily financed by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, under contract no. DE-AC02-05-CH11231 (D2S2 programme, KCD2S2) and the Laboratory Directed Research and Development Program of Lawrence Berkeley National Laboratory. The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under contract no. DE-AC02-05CH11231.

REFERENCES

(1) Grover, V.; Mandal, B. P.; Tyagi, A. K. *Handbook on Synthesis Strategies for Advanced Materials: Vol.-I: Techniques and Fundamentals*; Tyagi, A. K., Ningthoujam, R. S., Eds.; Springer: Singapore, 2021; pp 1−49, DOI: .

(2) Nwachukwu, I. M.; Nwanya, A. C.; Ekwealor, A. B. C.; Ezema, F. I. Research progress in solid-state synthesized $LiMnPO_4$ cathode material for Li-ion battery applications. *Appl. Surf. Sci. Adv.* **2023**, *18*, 100505.

(3) Kang, H.-C.; Jun, D.-K.; Jin, B.; Jin, E. M.; Park, K.-H.; Gu, H.-B.; Kim, K.-W. Optimized solid-state synthesis of $LiFePO_4$ cathode materials using ball-milling. *J. Power Sources* **2008**, *179*, 340−346.

(4) Buscaglia, M. T.; Bassoli, M.; Buscaglia, V.; Vormberg, R. Solid-State Synthesis of Nanocrystalline BaTiO3: Reaction Kinetics and Powder Properties. *J. Am. Ceram. Soc.* **2008**, *91*, 2862−2869.

(5) Todd, P. K.; Smith, A. M. M.; Neilson, J. R. Yttrium Manganese Oxide Phase Stability and Selectivity Using Lithium Carbonate Assisted Metathesis Reactions. *Inorg. Chem.* **2019**, *58*, 15166−15174.

(6) Han, H.; Lee, J. H.; Jang, H. M. Low-Temperature Solid-State Synthesis of High-Purity BiFeO3 Ceramic for Ferroic Thin-Film Deposition. *Inorg. Chem.* **2017**, *56*, 11911−11916.

(7) Onar, K.; Yakinci, M. E. Solid state synthesis and characterization of bulk {\ss} -FeSe superconductors. *J. Alloys Compd.* **2015**, *620*, 210−216.

(8) 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. *Adv. Mater.* **2021**, *33*, 2100312.

(9) Shi, L.; Gu, Y.; Qian, T.; Li, X.; Chen, L.; Yang, Z.; Ma, J.; Qian, Y. Synthesis of ultrafine superconducting MgB_2 by a convenient solid-state reaction route. *Phys. C* **2004**, *405*, 271−274.

(10) Merkle, R.; Maier, J. On the Tammann−Rule. *Z. Anorg. Allg. Chem.* **2005**, *631*, 1163−1166.

(11) Garn, P. D.; Habash, T. S. Solid state reactions. A rigorous test of the Hedvall effect. *J. Phys. Chem.* **1979**, *83*, 229−231.

(12) Szymkuć, S.; Gajewska, E. P.; Klucznik, T.; Molga, K.; Dittwald, P.; Startek, M.; Bajczyk, M.; Grzybowski, B. A. Computer-Assisted Synthetic Planning: The End of the Beginning. *Angew. Chem., Int. Ed.* **2016**, *55*, 5904−5937.

(13) Jain, A.; Hautier, G.; Moore, C. J.; Ping Ong, S.; Fischer, C. C.; Mueller, T.; Persson, K. A.; Ceder, G. A high-throughput infrastructure for density functional theory calculations. *Comput. Mater. Sci.* **2011**, *50*, 2295−2310.

(14) Doerr, S.; Harvey, M. J.; Noé, F.; De Fabritiis, G. HTMD: High-Throughput Molecular Dynamics for Molecular Discovery. *J. Chem. Theory Comput.* **2016**, *12*, 1845−1852.

(15) 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 Mater.* **2013**, *1*, 011002.

(16) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. *npj Comput. Mater.* **2015**, *1*, 1−15.

(17) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD). *JOM* **2013**, *65*, 1501−1509.

(18) Curtarolo, S.; Setyawan, W.; Hart, G. L. W.; Jahnatek, M.; Chepulskii, R. V.; Taylor, R. H.; Wang, S.; Xue, J.; Yang, K.; Levy, O.; Mehl, M. J.; Stokes, H. T.; Demchenko, D. O.; Morgan, D. AFLOW: An automatic framework for high-throughput materials discovery. *Comput. Mater. Sci.* **2012**, *58*, 218−226.

(19) Aykol, M.; Dwaraknath, S. S.; Sun, W.; Persson, K. A. Thermodynamic limit for synthesis of metastable inorganic materials. *Sci. Adv.* **2018**, *4*, No. eaaq0148.

(20) Sun, W.; Dacek, S. T.; Ong, S. P.; Hautier, G.; Jain, A.; Richards, W. D.; Gamst, A. C.; Persson, K. A.; Ceder, G. The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2016**, *2*, No. e1600225.

(21) Aykol, M.; Montoya, J. H.; Hummelshøj, J. Rational Solid-State Synthesis Routes for Inorganic Materials. *J. Am. Chem. Soc.* **2021**, *143*, 9244−9259.

(22) McDermott, M. J.; McBride, B. C.; Regier, C. E.; Tran, G. T.; Chen, Y.; Corrao, A. A.; Gallant, M. C.; Kamm, G. E.; Bartel, C. J.; Chapman, K. W.; Khalifah, P. G.; Ceder, G.; Neilson, J. R.; Persson, K. A. Assessing Thermodynamic Selectivity of Solid-State Reactions for the Predictive Synthesis of Inorganic Materials. *ACS Cent. Sci.* **2023**, *9*, 1957−1975.

(23) He, T.; Sun, W.; Huo, H.; Kononova, O.; Rong, Z.; Tshitoyan, V.; Botari, T.; Ceder, G. Similarity of Precursors in Solid-State Synthesis as Text-Mined from Scientific Literature. *Chem. Mater.* **2020**, *32*, 7861−7873.
(24) Kononova, O.; Huo, H.; He, T.; Rong, Z.; Botari, T.; Sun, W.; Tshitoyan, V.; Ceder, G. Text-mined dataset of inorganic materials synthesis recipes. *Sci. Data* **2019**, *6*, 203.
(25) Kim, E.; Huang, K.; Saunders, A.; McCallum, A.; Ceder, G.; Olivetti, E. Materials Synthesis Insights from Scientific Literature via Text Extraction and Machine Learning. *Chem. Mater.* **2017**, *29*, 9436−9444.
(26) McDermott, M. J.; Dwaraknath, S. S.; Persson, K. A. A graph-based network for predicting chemical reaction pathways in solid-state materials synthesis. *Nat. Commun.* **2021**, *12*, 3097.
(27) Szymanski, N. J.; Rendy, B.; Fei, Y.; Kumar, R. E.; He, T.; Milsted, D.; McDermott, M. J.; Gallant, M.; Cubuk, E. D.; Merchant, A.; et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* **2023**, *624*, 86−91.
(28) 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. *Nat. Synth.* **2024**, *3*, 606−614.
(29) Burger, B.; Maffettone, P. M.; Gusev, V. V.; Aitchison, C. M.; Bai, Y.; Wang, X.; Li, X.; Alston, B. M.; Li, B.; Clowes, R.; Rankin, N.; Harris, B.; Sprick, R. S.; Cooper, A. I. A mobile robotic chemist. *Nature* **2020**, *583*, 237−241.
(30) Szymanski, N. J.; Nevatia, P.; Bartel, C. J.; Zeng, Y.; Ceder, G. Autonomous and dynamic precursor selection for solid-state materials synthesis. *Nat. Commun.* **2023**, *14*, 6956.
(31) Pechukas, P. Transition State Theory. *Annu. Rev. Phys. Chem.* **1981**, *32*, 159−177.
(32) Harms, N.; Underkoffler, C.; West, R. Advances in Automated Transition State Theory Calculations: Improvements on the AutoTST Framework. *ChemRxiv*, 2020, https://chemrxiv.org/engage/chemrxiv/article-details/60c752ec9abda24460f8dec2 (accessed March 14, 2024).
(33) Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **1977**, *81*, 2340−2361.
(34) Andersen, M.; Panosetti, C.; Reuter, K. A Practical Guide to Surface Kinetic Monte Carlo Simulations. *Front. Chem.* **2019**, *7*, 202.
(35) Huang, G.; Montiel, D.; McAuliffe, R. D.; Veith, G. M.; Thornton, K. Phase-field modeling of solid-state metathesis reactions with the charge neutrality constraint. *Comput. Mater. Sci.* **2023**, *221*, 112080.
(36) Gardner, M. Mathematical Games. *Sci. Am.* **1970**, *223*, 120−123.
(37) Menshutina, N. V.; Kolnoochenko, A. V.; Lebedev, E. A. Cellular Automata in Chemistry and Chemical Engineering. *Annu. Rev. Chem. Biomol. Eng.* **2020**, *11*, 87−108.
(38) Sieradzki, L.; Madej, L. A perceptive comparison of the cellular automata and Monte Carlo techniques in application to static recrystallization modeling in polycrystalline materials. *Comput. Mater. Sci.* **2013**, *67*, 156−173.
(39) 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. *Nat. Commun.* **2018**, *9*, 4168.
(40) 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. *Nat. Mater.* **2020**, *19*, 1088−1095.
(41) Chase, M. *NIST-JANAF Thermochemical Tables*, 4th ed.; American Institute of Physics, −1, 1998.
(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. *Proc. Natl. Acad. Sci. U.S.A.* **2022**, *119*, No. e2209630119.
(43) Fatès, N. A Guided Tour of Asynchronous Cellular Automata. In *Cellular Automata and Discrete Complex Systems*; Heidelberg: Berlin, 2013; pp 15−30..
(44) Gallant, M. C.; Persson, K. A. pylattica: a package for prototyping lattice models in chemistry and materials science. *J. Open Source Softw.* **2024**, *9*, 6170.
(45) Beauger, A.; Mutin, J. C.; Niepce, J. C. Synthesis reaction of metatitanate $BaTiO_3$. *J. Mater. Sci.* **1983**, *18*, 3543−3550.
(46) Rom, C. L.; et al. Mechanistically Guided Materials Chemistry: Synthesis of Ternary Nitrides, $CaZrN_2$ and $CaHfN_2$. *J. Am. Chem. Soc.* **2024**, *146*, 4001−4012.
(47) 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. *J. Am. Chem. Soc.* **2021**, *143*, 15185−15194.
(48) Todd, P. K.; et al. Defect-Accommodating Intermediates Yield Selective Low-Temperature Synthesis of $YMnO_3$ Polymorphs. *Inorg. Chem.* **2020**, *59*, 13639−13650.
(49) Dosaev, K.; Istomin, S. Y.; Antipov, E. V. Synthesis and Crystal Structure of New Oxochloride $(Mn,Mg)_8Cl_3O_{10}$. *Russ. J. Inorg. Chem.* **2022**, *67*, 935−939.
(50) Ganose, A. et al. atomate2. 2024; https://github.com/materialsproject/atomate2 (accessed June 20, 2024).
(51) Adolphson, D. G.; Corbett, J. D. Crystal structure of zirconium monochloride. A novel phase containing metal-metal bonded sheets. *Inorg. Chem.* **1976**, *15*, 1820−1823.
(52) Imoto, H.; Corbett, J. D.; Cisar, A. Synthesis by hydrogen-driven disproportionation reactions. Synthesis and structure of the hexazirconium dodecahalide clusters Zr_6Cl_{12} and Zr_6Br_{12} and the double salt $Zr_6Cl_{12}.M_2ZrCl_6$ (M = sodium, potassium, cesium). *Inorg. Chem.* **1981**, *20*, 145−151.
(53) Schneider, S. B.; Frankovsky, R.; Schnick, W. Synthesis of Alkaline Earth Diazenides $MAEN_2$ (MAE = Ca, Sr, Ba) by Controlled Thermal Decomposition of Azides under High Pressure. *Inorg. Chem.* **2012**, *51*, 2366−2373.
(54) Krischner, H.; Kelz, G. Kristallstrukturbestimmung von Strontiumazid und Calciumazid. *Z. Anorg. Allg. Chem.* **1982**, *494*, 203−206.
(55) Reckeweg, O.; DiSalvo, F. J. Alkaline earth metal nitride compounds with the composition M_2N_X (M = Ca, Sr, Ba; X = H, Cl or Br). *Solid State Sci.* **2002**, *4*, 575−584.
(56) Taniguchi, T.; Dzivenko, D.; Riedel, R.; Chauveau, T.; Zerr, A. Synthesis of cubic zirconium(IV) nitride, c-Zr_3N_4, in the 6−8 GPa pressure region. *Ceram. Int.* **2019**, *45*, 20028−20032.
(57) Zerr, A.; Miehe, G.; Riedel, R. Synthesis of cubic zirconium and hafnium nitride having Th3P4 structure. *Nat. Mater.* **2003**, *2*, 185−189.
(58) Lerch, M.; Füglein, E.; Wrba, J. Synthesis Crystal Structure, and High Temperature Behavior of Zr_3N_4. *Z. Anorg. Allg. Chem.* **1996**, *622*, 367−372.

https://doi.org/10.1021/acs.chemmater.4c02301
Chem. Mater. XXXX, XXX, XXX−XXX