Lawrence Berkeley National Laboratory

Title

Combining Text Mining, in situ Characterization and ab initio Calculations to Rationalize BiFeO 3 Crystallization Pathways

Permalink

https://escholarship.org/uc/item/621982z6

Author

Sutter-Fella, Carolin

Combining text mining, in situ characterization and *ab initio* calculations to rationalize BiFeO₃ crystallization pathways

Maged Abdelsamie,[1,2,3‡], Kootak Hong [4,5‡], Kevin Cruse [1,6], Christopher J. Bartel[1,6,7], Viktoriia Baibakova [1,6], Amalie Trewartha [1,6,8], Anubhav Jain[9], Gerbrand Ceder [1,6], Carolin M. Sutter-Fella[10,]*

[1] *Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

[2] *current affiliation: Materials Science and Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia*

[3] *current affiliation: Interdisciplinary Research Center for Intelligent Manufacturing and Robotics, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia*

[4] *Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

[5] *Department of Materials Science and Engineering, Chonnam National University, Gwangju 61186, Republic of Korea*

[6] *Department of Materials Science and Engineering, University of California, Berkeley, CA 94720, USA*

[7] *current affiliation: Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, USA*

[8] *current affiliation: Toyota Research Institute, Los Altos, CA 94022, USA*

[9] *Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

[10] *Molecular Foundry Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

[‡] *These authors contributed equally to this work*
*Lead contact e-mail: csutterfella@lbl.gov

Keywords: Bismuth ferrite, crystallization pathway, in situ X-ray diffraction, text mining, ferroelectric materials, chelating agents, sol-gel

1

processing

Abstract: The combination of three highly complementary scientific domains is demonstrated to rationalize bismuth ferrite (BiFeO$_3$) [BFO] crystallization pathways: text mining to extract processing recipes from existing literature, in situ X-ray scattering to follow crystallization pathways of solution processed thin films, and *ab initio* calculations to develop a comprehensive understanding of thin-film formation from thermodynamic principles. Multiferroic BFO is chosen as an example material of interest for technological applications to demonstrate and validate this combined approach. Thermodynamic modeling showed that the production of a short-lived intermediate bismutite phase (Bi$_2$O$_2$CO$_3$) influences the energies of competing reactions forming BFO and Bi$_2$Fe$_4$O$_9$, increasing the thermodynamic driving force towards the formation of BFO relative to Bi$_2$Fe$_4$O$_9$ impurity phase over a large temperature window. The synergy between the scientific domains is exemplified through the rational guidelines developed for controlling high quality and phase pure material fabrication.

1. Introduction

The advancement of materials research requires a renewed focus on accelerating materials development and discovery, moving away from a reliance on empirical studies.[1] To achieve this goal, a paradigm shift is needed in which computational analysis will play an important role. There is now the opportunity to contribute to the early development and implementation of new integrated experimental, computational, and data informatics tools.[1,2] One complex problem of high interest requiring the integration of complementary scientific domains involves text mining to design thin film fabrication experiments, followed by in situ characterization of that fabrication, and the rational explanation of observed crystallization pathways.

2

The material system chosen here to develop and validate this approach is bismuth ferrite ($BiFeO_3$, BFO). BFO was chosen based on a cursory search over our literature database for oxide papers discussing impurity phases appearing in fully processed thin films, as it was the most data-rich having the highest number of papers returned (followed by $SrTiO_3$ and $LiFePO_4$) and can be processed from solutions. BFO has gained attention as a room-temperature multiferroic,[3,4] is environmentally friendly (as compared to lead-based $Pb(Zr,Ti)O_3$ material), and is considered one of the most promising materials in next-generation non-toxic ferroelectric memory and spintronic devices.[5] BFO exhibits a rhombohedral distorted perovskite crystal structure with the space group R3c at room temperature.[6,7] Despite its promise, the formation of impurity phases in synthesizing BFO significantly jeopardizes its ferroelectric properties due to the leakage currents in the film. This detriment makes it essential to synthesize phase-pure polycrystalline BFO films. Bismuth-rich ($Bi_{25}FeO_{40}$) or iron-rich ($Bi_2Fe_4O_9$) phases are the most commonly observed impurities as their formation is thermodynamically competitive with BFO formation when starting from typical precursors. In order to control phase purity via targeted processing and precursor choice, an understanding of the crystallization pathway is necessary, but studies on the crystallization pathway of BFO thin films are rare.[8] It is noted that impurity phases are not the only metric to define film quality but was chosen here as a starting point for methodology development.[9]

Thin films are the most common material form for device applications in the semiconductor industry.[10,11] Among various techniques to fabricate thin films (e.g., RF-sputtering, molecular beam epitaxy, pulsed laser deposition, magnetron sputtering, chemical solution deposition, chemical vapor deposition, and sol-gel processing),[11-17] sol-gel processing is often prioritized due to lower processing temperatures and more precise compositional control, allowing the fabrication of cost-effective, large-area, and high-purity thin films.[10,11] In sol-gel processing, crystallization pathways, final film microstructure, and properties can be tuned by

3

solution engineering (e.g., using pure or mixed solvents, adding chelating agents)[18-20] and experimental conditions (e.g., temperature, time, number of iterations), but the causal relationship has been observed only partially due to limitations of conventional approaches.

To establish the link between sol-gel processing parameters and BFO crystallization pathways, we developed and validated an integrated approach combining text mining, in situ X-ray scattering, and *ab initio* calculations (**Figure 1**). First, we used text mining to search existing research publications and create structured datasets to design efficient experiments based on available scientific knowledge.[9] Next, we conducted in situ X-ray diffraction (XRD) experiments to follow the crystallization pathway, observe the formation of impurity phases, and compare these against reported impurity phases extracted from literature. Last, we used density functional theory calculations to rationalize observed crystallization pathways.
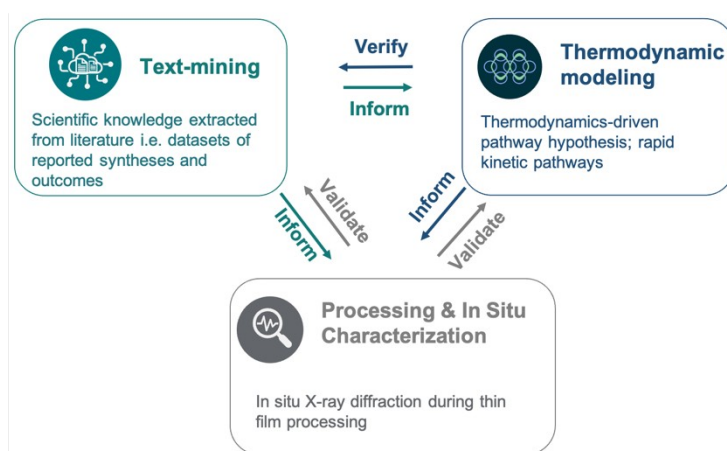

Figure 1 Schematic representation of the interplay between text mining, processing & in situ characterization, and thermodynamic modeling using *ab initio* calculations.

In summary, text mining indicated the prevalence of experimental decisions for various processing parameters as well as the correlation between processing temperatures and impurity phase formation. In situ XRD revealed the metastable formation of the intermediate bismutite phase ($Bi_2O_2CO_3$). *Ab initio* calculations found that the presence of this

$Bi_2O_2CO_3$ phase increases the thermodynamic driving force to form BFO over $Bi_2Fe_4O_9$ over a large temperature window. It is noted that impurity phases are not the only metric to define film quality but was chosen here as a starting point for methodology development.[9] We use the information provided by the three techniques to design efficient experiments, refine our understanding of the BFO crystallization pathway, and provide rational guidelines for precise control of BFO crystallization.

2. Results and discussion:

2.1 Text mining of BFO fabrication

Text mining was employed in this study to provide a comprehensive picture of the existing variable space explored for sol-gel processed BFO thin films. This helped to inform experiments and in validating mechanistic insights found through in-situ XRD characterization and *ab initio* calculations. Through the manual text mining of existing literature on the sol-gel fabrication of BFO thin films, we compiled and analyzed a dataset of 340 thin film processing procedures and outcomes from 178 publications. Details for the collection and extraction of these processing recipes are provided in Section 4 (*Article identification for text-mined dataset and Processing procedure and outcome extraction*) as well as in an accompanying manuscript. [Kevin Cruse et al. in preparation] This dataset was constructed via a search over an in-house literature database as well as Clarivate Analytics' Web of Science for papers relevant to BFO fabrication, followed by manual inspection to determine which papers contained fabrication of undoped BFO thin films through the sol-gel method. Although key information from a limited number of publications can be extracted by humans, the dataset was manually parsed and annotated to pursue two goals: first, to create a 'gold' standard to train machine learning models that predict phase purity from processing conditions, and second, to guide the planning of experimentalists (K. Cruse *et al*., manuscript in preparation). For the latter, with data analysis, we revealed main thin film fabrication steps, common processing values,

5

underexplored processing variables, and designed new experiments. Thus, text-mining provided the opportunity to process a notably bigger scope of data than is commonly treated by a human researcher, increasing the confidence about derived conclusions and allowing us to study all aspects of BFO sol-gel fabrication in detail.

A typical sol-gel fabrication of BFO thin films is schematically depicted in Figure **S1** and involves the liquid precursor preparation, followed by spin-coating, baking, and annealing. **Figure 2** illustrates a visual summary of the different aspects of the BFO processing space as reflected in this dataset, providing a comprehensive overview of typically reported impurity phases and their frequency, solvents, chelating agents, and mixtures thereof used, and lastly temperatures used during different processing stages. While various metal precursors have been used for BFO fabrication (**Figure 2**a), text mining shows that iron nitrate nonahydrate [$Fe(NO_3)_3 \cdot 9H_2O$] combined with bismuth nitrate pentahydrate [$Bi(NO_3)_3 \cdot 5H_2O$] comprise the overwhelming majority of recipes (318 in total, 95% of the dataset; extreme is not included in heat colors in **Figure 2**a to emphasize differences in lower proportion combinations). Although Bi and Fe are stoichiometric in the $BiFeO_3$ crystal structure, different Bi:Fe ratios in the precursor solution were reported in the literature (**Figure S2**a). A Bi:Fe ratio of 1.05:1 was the most frequently used, mainly as an attempt to compensate for Bi volatility during annealing. A wide range of solvents (**Figure S2**b) has been employed to synthesize BFO thin films using the sol-gel method, including 2-methoxyethanol (2ME), ethylene glycol (EG), 2-ethoxyethanol, ethanol, dimethylformamide, and propionic acid. Text mining reveals that 2ME and EG with 248 and 72 cases, respectively, and their mixtures (19 cases) are most frequently used to fabricate BFO films. The use of chelating agents has been adopted to increase the viscosity and drying time of the gel in order to obtain high-quality and uniform films. Several chelating agents, including acetic acid, citric acid, nitric acid, acetylacetone, and ammonium hydroxide, were used, with acetic and citric acids combined with other solvents (114 and

6

60 cases) having the largest share (Venn diagrams **Figure 2**b-c, and **S2**c). 2ME is most frequently combined with acetic acid (84 cases) or citric acid (34 cases). Dehydrating agents and surfactants, such as acetic anhydride and ethanolamine, are also common reagents in this processing space, though they are not included in this analysis. There are only a few papers that study the effects of varying the combination of precursor solution agents across experiments.[21,22] Qi *et al*., found $Bi_2Fe_4O_9$ in the acetic route (i.e. acetate-based solutions using bismuth(III) acetate, iron(III) acetylacetonate, acetic acid precursors) after film sintering and speculated that the Bi evaporation rate is higher for the acetyl solution. While no impurity phases were observed in the citric route (i.e. citrate-based solutions using Bi_2O_3, nitric acid, iron(III) citrate precursors).[21] The study by Liang *et al*. reported no impurity phases for both, nitric acid and mixed 2ME and glacial acetic acid solutions.[22] To this end, little is known about the effects of solvent engineering solely or combined with chelating agents on the BFO crystallization pathway. Text mining showed that roughly a quarter of the reported syntheses result in phase impurities as depicted in **Figure 2**d together with the most common impurity phases and their formulae. $Bi_2Fe_4O_9$, binary oxides Bi_2O_3 and Fe_2O_3, and bismuth-rich impurities make up a large portion of the possible impurity phases formed. **Figure S2**d shows the prevalence of annealing environments, with air being the most common. The oxygen atmosphere is also used on occasion to mitigate the development of oxygen vacancies during processing.

According to our text mining analysis, the annealing temperature is another important processing variable. The following describes the purpose of various heat treatment steps throughout a typical BFO fabrication (see schematic **Figure S1**). After spin-coating, individual wet films are dried at ~100-250 °C to evaporate the solvents and induce gelation in the as-spin-coated films. Then, the films are pre-baked at ~200-400 °C to induce the thermal decomposition of organic materials. Up to this stage, the films are mostly amorphous. Lastly, the films are

7

annealed at relatively higher temperatures of ~400-800 °C to induce nucleation, growth, and conversion of the amorphous to a polycrystalline film. These steps are typically administered after each of several layers are spin-coated ("layer-by-layer") until the desired film thickness is reached, after which a final annealing step is often employed. Using text mining, we studied the distribution of layer-by-layer and final annealing temperatures and the relationship between temperature and impurity phases, which is summarized in **Figure 2**e. The typical crystallization window for sol-gel-derived BFO thin films is found to be between ~500 °C and ~600 °C. There is a higher probability of impurity phase formation with a relatively higher temperature than the average annealing temperature. For Instance, processing procedures with annealing at the most common temperature (~ 550 °C) result in impurities in 6% and 14% of cases for layer-by-layer and final annealing, respectively. With the increase of annealing temperatures to 600 °C, the proportion of syntheses that lead to impurities changes to 4% and 36% for layer-by-layer and final annealing, respectively. At 700 °C, this proportion increases to 100% for layer-by-layer (our dataset has records of only phase-impure films when annealing at this temperature) and 80% for final annealing.
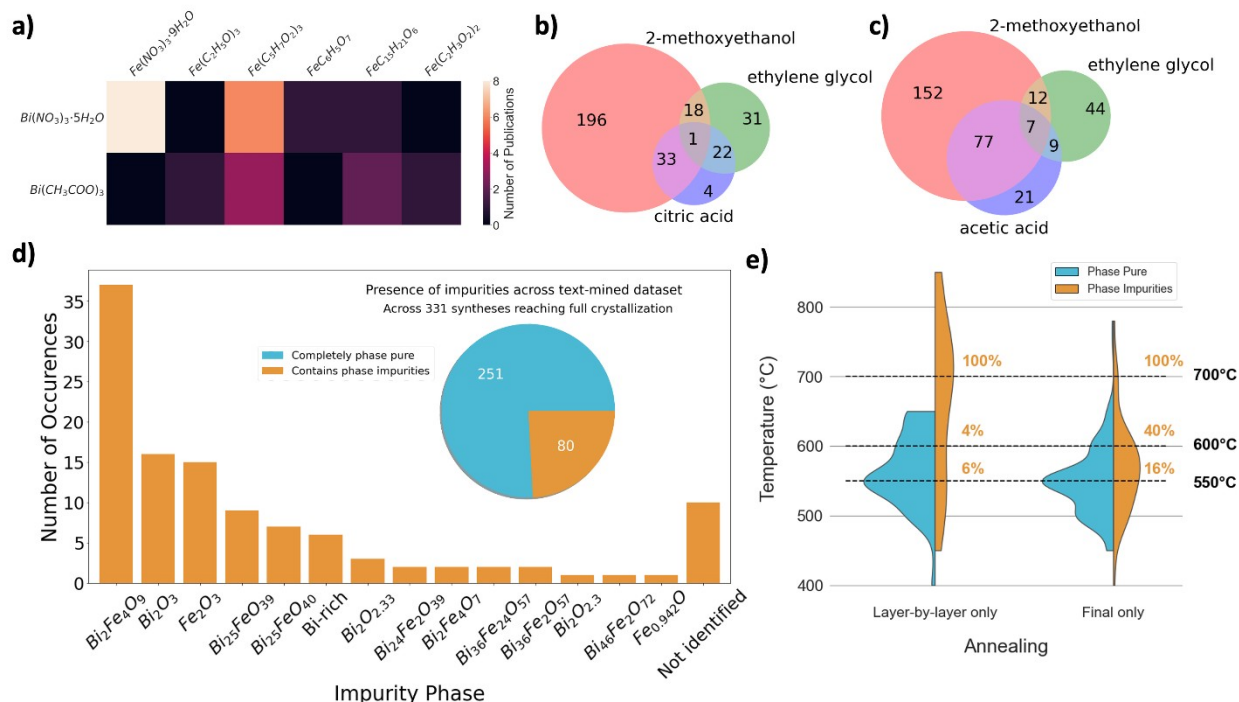
Figure 2 Overview of text-mined experimental conditions employed in BiFeO₃ thin film solution syntheses published in the literature. (a) Heatmap depicting the frequency of combinations of Bi and Fe precursors. Iron precursors from left to right (without metal prefix) are iron: nitrate nonahydrate, pentethoxide, acetylacetone, citrate, pentanedionate, and acetate. Bismuth precursors from top to bottom are bismuth: nitrate pentahydrate and acetate. (b-c) Count of recipes with the most common solvents 2-methoxyethanol (2ME) and ethylene glycol (EG) and their combination with the chelating agents citric acid in and acetic acid in (b) and (c), respectively. (d) Prevalence of the formation of impurity phases across all syntheses in this dataset, along with the formulae for the most common impurity phases that form out of these syntheses. (e) Distribution of temperatures used during film annealing including information on presence/absence of impurity phases. Percentages represent the proportion of phase impure syntheses to phase pure within a 20 °C temperature window below and above the dotted line.

2.2 In situ X-ray diffraction during BFO annealing:

With the input from text mining to design efficient experiments for the BFO film fabrication, we investigated the role of solvents, chelating agents, and metal ratios on the crystallization pathway via in situ XRD

9

during thermal annealing. The fabrication of BFO films is described in detail in the experimental section. In brief, $Bi(NO_3)_3 \bullet 5H_2O$ and $Fe(NO_3)_3 \bullet 9H_2O$ precursor salts were dissolved in 2ME, EG, or mixtures thereof. Some recipes include chelating agents (citric acid or acetic acid). The final film is built from 5 cycles of spin coating and pre-baking at 300 °C. **Figure 4** and **S1** show the summary of the evolving phases, and **Figures S3** and **S4** provide corresponding 2D XRD color maps. Due to a compromise between data acquisition time and the lifetime of possible intermediate phases, a relatively small 2theta range from 10-35° was chosen. Prior to in situ diffraction measurements ex situ diffraction measurements were performed on a wider 2theta range to more confidently evaluate impurity phase formation (FigureS**5**). The majority of observed results fall into three different categories: (i) no impurity phase, (ii) persistent impurity phase, and (iii) intermediate phase (**Figure 3**). Line profiles of these measurements with peak labeling are illustrated in **Figure S6**. The temperature profile and holding times during in situ measurements are depicted in **Figure S7**. The 2D color maps (Figure 3) show selected experiments with Bi:Fe ratio of 1.1, as is widely recommended in the literature to compensate for Bi loss during annealing, and varied precursor solutions: pure 2ME, mixed solvents EG:2ME (9:1), and 2ME with citric acid. Figures S3 and S4 illustrate a broader range of experiments with Bi:Fe ratios of 0.9-1.1 and precursor solutions of pure EG and acetic acid additive. Category (i), no impurity phase, corresponds to solvent systems utilizing pure 2ME or EG. The precursor does not exhibit crystalline impurity phases throughout the annealing process, directly transforming to phase pure BFO upon reaching a certain temperature inversely proportional to Bi content (Figure S3a-f). In category (ii), an impurity phase Bi_2O_3 was observed at the pre-bake stage for mixed solvent systems EG:2ME. The Bi_2O_3 phase persists onward and dissociates during thermal annealing at a temperature directly proportional to Bi content, as revealed by the disappearance of Bi_2O_3 diffraction peaks

10

followed by the onset of BFO crystallization (Figure S3g-l). Lastly, in category (iii), an intermediate phase was observed for samples prepared with the addition of citric acid to 2ME. Although the pre-baked films are amorphous, an intermediate phase, bismutite or bismuth subcarbonate ($Bi_2O_2CO_3$), appears with the initial crystallization of the BFO phase. $Bi_2O_2CO_3$ forms the earliest for Bi-rich (Bi:Fe = 1.1) samples (475 °C) and dissociates at 525 °C. For both stoichiometric (Bi:Fe = 1.0) and Bi-deficient (Bi:Fe = 0.9) samples, the $Bi_2O_2CO_3$ phase forms upon reaching 500 °C and dissociates at 550 °C (Figure S4d-f).

The growth of nuclei in BFO films was reported to occur via oriented attachment, meaning the coalescence of similarly oriented crystals.[8] Our in situ results show that the bismutite phase appearance correlates with the onset of BFO crystallization (Figure S3). Several studies discussed the occurrence and role of intermediate carbonate phases in perovskite-type oxides.[23,24] For example, it was previously suggested that in BFO processing with acetic acid, bismuth acetate, iron acetylacetonate, and water, $Bi_2O_2CO_3$ crystallizes first in the amorphous film at the pre-existing metal-organic clusters before BFO formation.[8] The combination of acetic acid with EG or 2ME or citric acid with EG in our study with bismuth and iron nitrates, however, did not result in the bismutite phase or any other intermediate phases (Figures S3, S4). Thus, we speculate that the carboxylic group in citric acid contributes to the formation of $Bi_2O_2CO_3$ in the presence of 2ME.

It was found in this study, that the onset of BFO crystallization is affected by the choice of solvents, chelating agents, and the metal ratio of the precursor solution. The overall trend for all conditions is that a lower onset crystallization temperature occurs with a higher Bi content in the precursor solution. For instance, stoichiometry change from Bi-deficient to Bi-excess lowers the onset of BFO crystallization by 25 - 50 °C. For films prepared using pure 2ME, the onset of BFO crystallization is reduced from 550 °C for Bi-deficient samples (Bi:Fe = 0.9) to 525 °C for both

stoichiometric (Bi:Fe = 1.0) and Bi-excess (Bi:Fe = 1.1) samples (**Figure 4**). Films prepared from pure EG versus pure 2ME show a lower BFO crystallization onset. For instance, for films prepared from pure EG, the onset of crystallization is reduced from 475 °C for Bi-deficient samples (Bi:Fe = 0.9) to 450 °C for both stoichiometric (Bi:Fe = 1.0) and Bi-excess (Bi:Fe = 1.1) samples (**Figure 4** and **Table S1**). Higher Bi content seems to facilitate BFO growth,[8] possibly related to the nature of the pre-existing metal-organic clusters and their transformation to the rhombohedral distorted perovskite structure. The lowest crystallization onset at 450 °C was found for mixed EG:2ME solvents (category ii) with Bi_2O_3 as a persistent impurity phase. Bi_2O_3 dissociates at a higher temperature, the higher the Bi content. Notably, the BFO crystallization onset is independent of the Bi:Fe ratio for EG:2ME = 7:3, exhibiting the lowest crystallization onset at 450 °C for all stoichiometric compositions. Adding chelating agents (acetic or citric acid) to pure solvents results in an increase in the onset of BFO crystallization compared to films prepared from pure solvents. We hypothesize that this increase can be attributed to the formation of amorphous complexes that require higher temperatures to dissociate before allowing BFO crystallization as compared to pure EG or 2ME samples, which may not be able to coordinate with the metals in solution strongly. Although EG has a higher boiling point than 2ME (197 °C as compared to 124 °C), samples prepared from EG with chelating agents still have a lower BFO onset crystallization temperature than samples prepared from 2ME with chelating agents. This behavior confirms the ability of EG to reduce the onset of BFO crystallization, as discussed earlier.

Text mining analysis revealed $Bi_2Fe_4O_9$ and Bi_2O_3 as the main impurity phases and a higher probability for impurity phase formation at higher annealing temperatures, which is somewhat confirmed by the in situ XRD experiments. On the other hand, the formation of $Bi_2O_2CO_3$ was observed as a short-lived intermediate phase during the in situ thermal annealing experiments, but this phase is not part of the final film and thus it was not

12

extracted via text mining. Finding this intermediate phase through in situ XRD demonstrates the importance of our investigative approach that integrates detailed, time-resolved phase identification for the fabrication of our system along with historical collection of processing data.
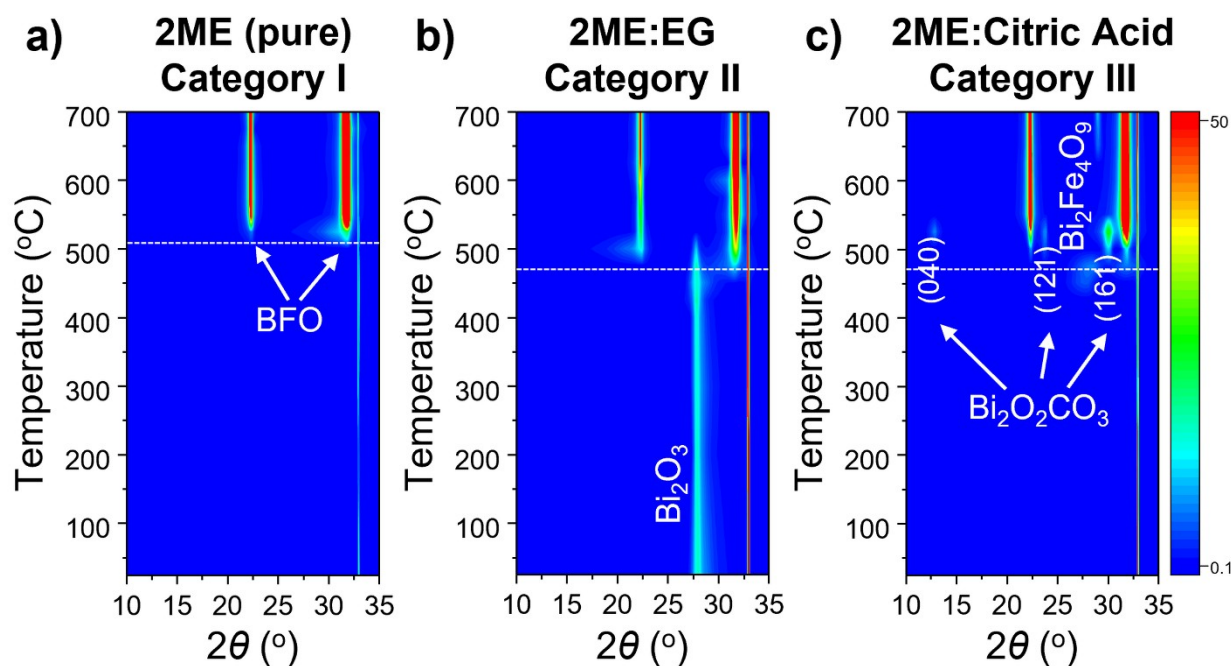

Figure 3 2D color maps of in situ X-ray diffraction during thermal annealing of BiFeO$_3$ thin films with chemistries representative of three categories (a) no impurity phase, (b) persistent impurity phase Bi$_2$O$_3$, and (c) intermediate phase Bi$_2$O$_2$CO$_3$. The films were spin coated using a precursor solution with stoichiometry Bi:Fe = 1.1 with pure 2ME (a), mixed solvents EG:2ME (b), and 2ME with citric acid (c). The dotted lines mark the BFO crystallization onset. The 'shadows' before the BFO peaks appear (a,b) as well as the broad peak ~475°C (c) are artifacts from the plotting since there are no visible peaks at these points in time at these angles (see Fig. S6).
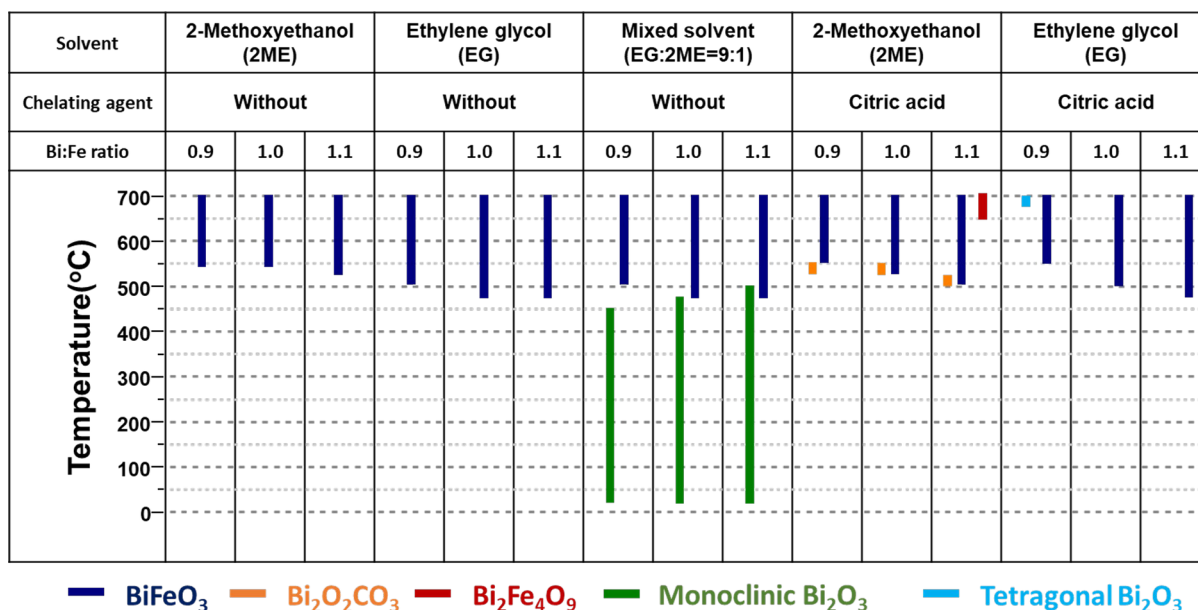
Solvent	2-Methoxyethanol (2ME)			Ethylene glycol (EG)			Mixed solvent (EG:2ME=9:1)			2-Methoxyethanol (2ME)			Ethylene glycol (EG)		
Chelating agent	Without			Without			Without			Citric acid			Citric acid		
Bi:Fe ratio	0.9	1.0	1.1	0.9	1.0	1.1	0.9	1.0	1.1	0.9	1.0	1.1	0.9	1.0	1.1

Figure 4 Summary of evolving phases during thermal annealing of BFO films prepared from different solvents

2.3 Thermodynamics of BFO vs impurity phases

As the last scientific domain deployed here, density functional theory (DFT) calculations with the meta-GGA SCAN density functional were used to assess the thermodynamic stability of BFO with respect to decomposition into binary and ternary competing phases.[25] A recently introduced model for finite temperature thermodynamics was used to map DFT-calculated formation energies at 0 K to temperature-dependent Gibbs formation energies, $\Delta G_f(T)$.[26] In **Figure 5**a, we show the calculated Bi-Fe-O phase diagram at 627 °C (900 K), where $BiFeO_3$ is calculated to be thermodynamically stable with respect to decomposition into $Bi_2Fe_4O_9$ and $Bi_{25}FeO_{39}$. By overlaying the thermodynamic driving force for compound formation from the elements, ΔG_f, on the ternary phase diagram, we see there may be a preference for forming Fe-rich phases as this region of the phase diagram has the largest driving force, indicated by the lighter

14

coloring.[27]

Both BFO and $Bi_2Fe_4O_9$ contain Bi^{3+} and Fe^{3+}, meaning in the Bi-Fe-O phase diagram, they lie along the tie-line formed by the corresponding binary phases having the same nominal oxidation states Bi_2O_3 and Fe_2O_3. In **Figure 5**b, we show that only BFO and not $Bi_2Fe_4O_9$ is thermodynamically stable (on the convex hull)[28] along this tie line at low temperature (27 °C, 300 K), but increasing the temperature to 627 °C / 900 K (**Figure 5**c), preferentially stabilizes $Bi_2Fe_4O_9$ compared to BFO. While at 300 K, BFO is the only stable ternary phase along this tie line, at 900 K, BFO and $Bi_2Fe_4O_9$ are both thermodynamically stable, and there is comparable thermodynamic driving force to form either phase from the binary oxides (Fe_2O_3, Bi_2O_3). This indicates the increasing possibility of Fe-rich impurity phase formation for syntheses where the sample is exposed to high temperature for extended periods. Indeed, $Bi_2Fe_4O_9$ is a common impurity phase formed during BFO fabrication as revealed by text mining results (Figure 2a) and shown in Figure 3c.[29] Moreover, increasing the temperature further preferentially stabilizes $Bi_2Fe_4O_9$ relative to BFO. This finding explains both the text mining and in situ XRD results (**Figure 4**), where both revealed that impurities are more likely to form at relatively higher temperatures.

Although text mining did not reveal $Bi_2O_2CO_3$ as an impurity phase occurring in fully fabricated BFO thin films, it was observed as a short-lived intermediate phase during in situ XRD measurements. Highlighting the effect of this intermediate on competing reactions in this system, our DFT calculations (**Figure 5d**) showed that the presence of $Bi_2O_2CO_3$ lowers the free energy of the reaction, ΔG_{rxn}, toward $BiFeO_3$ as compared to the reaction toward $Bi_2Fe_4O_9$ (on a per atom basis).[30] That is, the relative driving force toward the desired phase, $BiFeO_3$, compared to the undesired impurity, $Bi_2Fe_4O_9$, increases when $Bi_2O_2CO_3$ is present as an intermediate, largely because more CO_2 is evolved per atom of reaction when BFO rather than $Bi_2Fe_4O_9$ is formed from Fe_2O_3 and $Bi_2O_2CO_3$. This confirms the influence of this intermediate on the crystallization pathway.
15

Figure 5d illustrates the change in free energy of reaction for BFO versus $Bi_2Fe_4O_9$ formation in the absence versus presence of the $Bi_2O_2CO_3$ phase.

The calculated reaction energies for reactions without $Bi_2O_2CO_3$, solid lines shown in Fig. 4d are similar and there is a crossover point at ~650°C where $Bi_2Fe_4O_9$ appears to be thermodynamically preferred over BFO as shown previously in **Figure 5b-c**. With more confidence one can extract from these calculations that $Bi_2Fe_4O_9$ is favored compared to BFO at higher temperatures. It is pointed out that the slopes of these lines are distinguishably different. In the presence of $Bi_2O_2CO_3$ phase, however, $BiFeO_3$ is thermodynamically preferred to $Bi_2Fe_4O_9$ across the entire temperature range. This finding reveals that the formation of the BFO phase becomes more favorable as compared to the Fe-rich impurity phase ($Bi_2Fe_4O_9$) in the presence of $Bi_2O_2CO_3$, and this finding is corroborated by the text-mined data. The experimental limit for the in situ measurements at 700°C almost coincides with the crossover point of the calculated reaction energies and the calculations may underestimate the crossover point. Lastly, BFO seems to be the kinetically preferred phase and if it starts forming at low temperatures where it is also thermodynamically preferred, formation of the $Bi_2Fe_4O_9$ phase is inhibited at higher temperatures due to a small driving force of $Bi_2O_3 + Fe_2O_3 \rightarrow Bi_2Fe_4O_9$. Consequently, the conversion rate of BFO to $Bi_2Fe_4O_9$ is slow. Experiments however show that not only the temperature and presence/absence of $Bi_2O_2CO_3$ influence impurity formation, the metal ratio also plays a role. With an intermediate $Bi_2O_2CO_3$ phase and under Bi-rich processing, $Bi_2Fe_4O_9$ forms at ~650 °C while this Fe-rich phase does not form under Bi-poor or stoichiometric processing conditions within the experimentally accessible temperature window (Figure S4d-f).

It should be noted that there are various sources of error that could influence the computed thermodynamics, including errors in the 0 K compound thermochemistry computed using DFT and in the vibrational entropy correction using the model presented in Ref. [26]. However, most of these errors in absolute formation energies are likely to cancel in the

16

calculation of relative energies due to systematic dependence of the errors on chemical composition. This is especially true when comparing differences in reaction energies, for instance as it pertains to the relative preference for BFO over $Bi_2Fe_4O_9$ when the carbonate precursor is used. **Figure 5**e shows the prevalence of the formation of impurity phases (extracted by text mining) across all fabrications using 2ME, 2ME with citric acid, and 2ME without citric acid. Text mining confirms that the use of 2ME with citric acid (associated with $Bi_2O_2CO_3$ formation as in our in-situ XRD results) has a probability of 100% of obtaining pure BFO films as compared to 2ME without citric acid (67.5%). Through text mining of historic syntheses we validate our findings (i.e. phase impurity formation is mitigated by the presence of $Bi_2O_2CO_3$ from the inclusion of citric acid) from experiment and theory because all published procedures using 2ME + CA in the precursor solution resulted in a phase pure outcome.
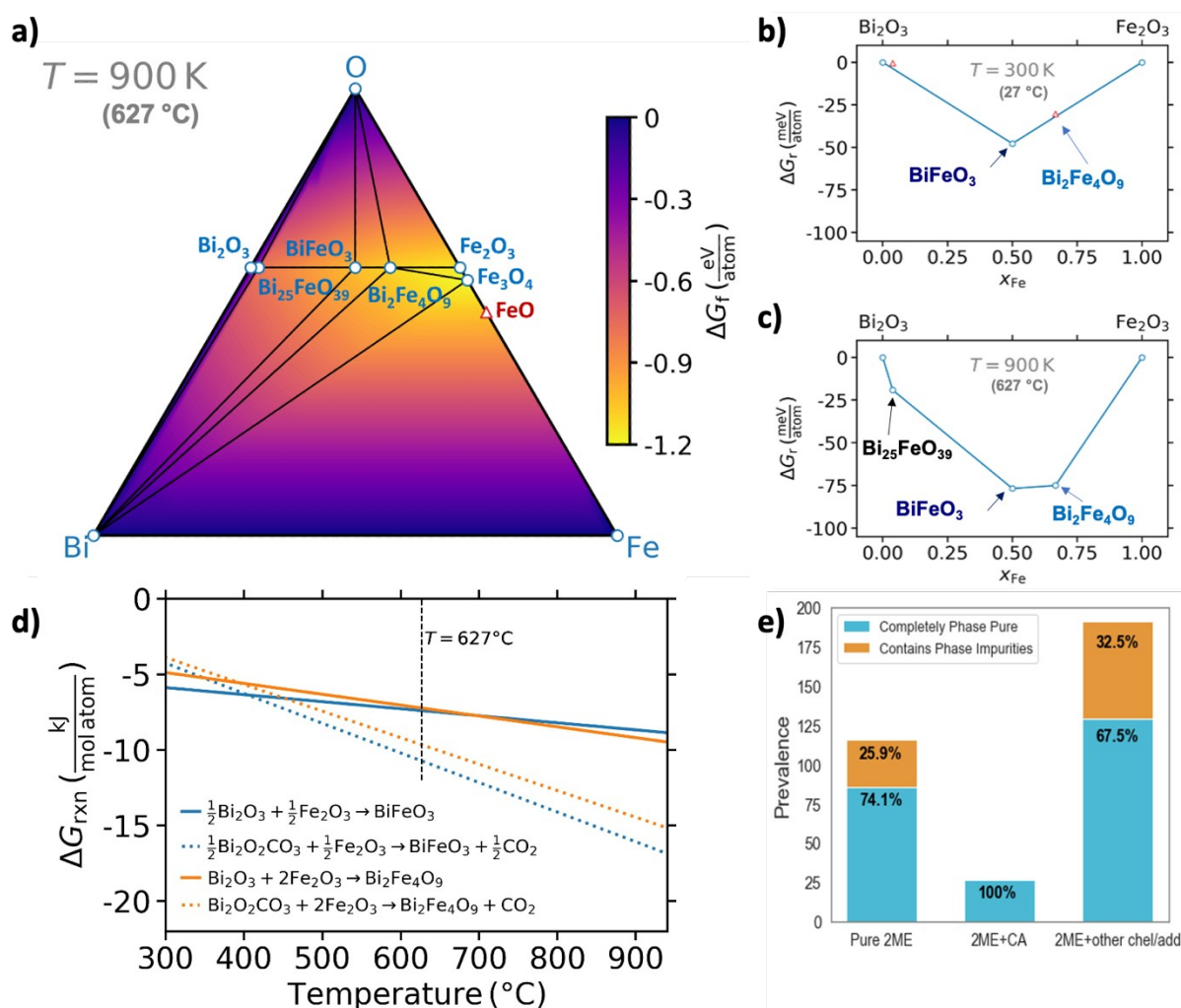
Figure 5 Calculated thermodynamics in the Bi-Fe-O chemical space. (a) Ternary Bi-Fe-O phase diagram at 627 °C (900 K). Blue circles indicate stable phases. The color-bar indicates the driving force for phase formation with respect to the elemental phases (Bi, Fe, O). At stable compositions (e.g., BiFeO$_3$), the colorbar indicates the formation energy of this phase. At points in composition space that do contain a stable compound, the formation energy (colorbar) is computed relative to the combination of stable phases that lie on the convex hull and minimize the formation energy at that composition (the decomposition products). (b-c) Driving forces for phase formation with respect to the Bi$_2$O$_3$-Fe$_2$O$_3$ tie line at (b) 27 °C (300 K) and at (c) 627 °C (900 K). (d) Reaction energies for the formation of BFO and Bi$_2$Fe$_4$O$_9$ from precursors without or with the Bi$_2$O$_2$CO$_3$ phase. For comparing reactions with diverse compositions, each reaction energy is normalized per mole of atoms. For clarity, the reactions in the figure legend are shown on the basis of 1 mole of ternary Bi-Fe-O phase formed. To compute the reaction energy using the kJ/(mol atom) basis, the energies for the molar reactions shown in the legend are divided by the number of atoms in the product side of each reaction (5 atoms for the reaction forming BFO, 6.5 for the reaction forming BFO and CO$_2$, 15 for the reaction forming Bi$_2$Fe$_4$O$_9$, and 18 for the reaction forming

$Bi_2Fe_4O_9$ and CO_2). (e) Prevalence of the formation of impurity phases across all syntheses using 2ME, 2ME with citric acid, and 2ME without citric acid normalized per each batch size. Completely pure films and films containing impurities are represented as blue and orange bars, respectively.

3. Discussion and Conclusions

Based on the complementary input from text mining, in situ XRD, and thermodynamic modeling, we propose a comprehensive understanding of BFO crystallization pathways. **Figure 6** shows an overview of the flow of information between the three domains. Text mining provided input on the most frequently used chemicals and processing conditions, identified the most commonly reported impurity phases ($Bi_2Fe_4O_9$ and Bi_2O_3), and revealed a higher probability of impurity phase formation at relatively higher annealing temperatures. Consequently, in situ XRD measurements were performed on the most commonly used precursor compositions and confirmed the formation of persistent impurity phases ($Bi_2Fe_4O_9$ and Bi_2O_3) at relatively higher annealing temperatures, validating text mining results. The latter was confirmed and explained by the *ab initio* calculations because the Gibbs formation energy of $Bi_2Fe_4O_9$ is close to BFO at high temperatures (900 K). Interestingly, in situ XRD measurements revealed a short-lived intermediate bismutite phase ($Bi_2O_2CO_3$) which was not identified through text mining. *Ab initio* calculations uncovered the role of this short-lived intermediate phase and predicted that BFO crystallization becomes more favorable than impurity phase formation in the presence of $Bi_2O_2CO_3$ by lowering the system's free energy. Text mining supports this finding and has shown that the use of 2ME with citric acid (associated with $Bi_2O_2CO_3$ formation as revealed by in situ XRD) has a higher probability of obtaining pure BFO films than other chemical combinations. Moreover, since both text mining and *ab initio* calculations predict that the formation of impurities is more likely at higher temperatures, in situ XRD was used to detect the parameters that can allow for lowering the annealing

19

temperature, such as the solvent choice (using pure EG or mixed EG:2ME) as well as the used of Bi-rich precursors.
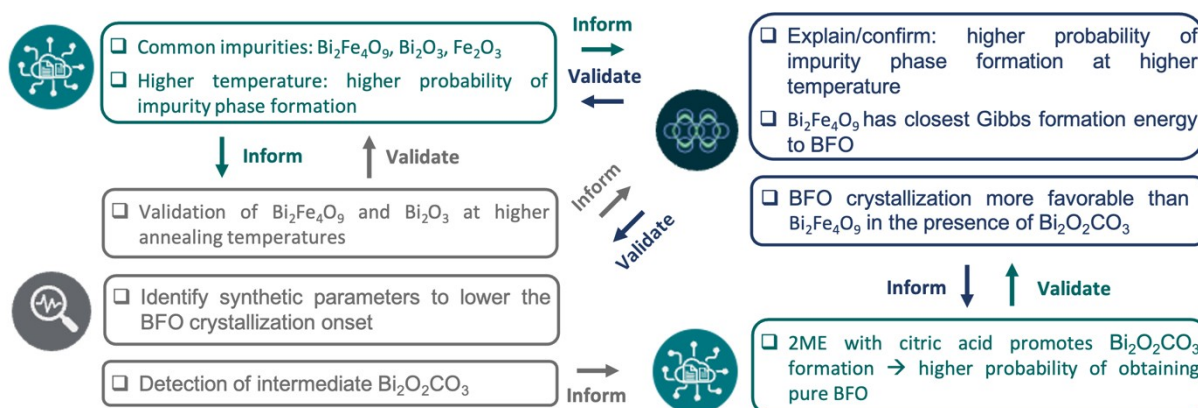

Figure 6 Schematic representing an overview of the flow of information between text mining, in situ XRD, and thermodynamic modeling.

In summary, combined text mining, in situ characterization, and thermodynamic modeling helped to gain a comprehensive understanding of the BFO crystallization pathway. In particular, we identified the thermodynamic origin of the intermediate $Bi_2O_2CO_3$ phase in favoring the formation of BFO instead of $Bi_2Fe_4O_9$. The choice of the solvent, 2ME in combination with citric acid as a chelating agent, promotes the formation of $Bi_2O_2CO_3$. In addition, the solvent impacts the BFO crystallization temperature, where our results reveal that EG effectively reduces the onset temperature whether it is used alone or mixed with 2ME, or in combination with chelating agents. Both text mining and *ab initio* calculations predicted that the formation of impurities is more likely at higher temperatures. Hence recipes that can allow for lowering the annealing temperature are favored and include solvent choice as well as Bi-rich precursors. This study illustrates how the combination of text mining, in situ characterization, and *ab initio* calculations helped to design relevant experiments, reveal intermediate and impurity phases, and understand their role and formation probability. With this comprehensive combination of scientific domains, processing recipes can be designed

efficiently to fabricate high-quality and phase-pure materials instead of following costly and time-consuming trial-and-error approaches. The methodologies applied here serve as stepping stones toward closed loop workflows.

4. Resource Availability

Lead Contact: *Carolin M. Sutter-Fella,* csutterfella@lbl.gov

Materials Availability: All materials used in this work are commercially available.

Data and Code Availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The dataset used for text mining is available on GitHub: https://github.com/kevcruse96/bfo-impurityphase-analysis

All data and code associated with this study are available from the lead contact upon request.

Article identification for text-mined dataset: The corpus of 178 articles for our dataset of sol-gel processing of $BiFeO_3$ thin films was compiled through a combination of automated text mining and manual extraction via the Web of Science. To begin our search, we used regular expression matching to identify articles from published datasets of text-mined inorganic processing procedures[31] from the years 2000-2020 containing terms such as "impurity", "secondary phase", "phase pure", etc. Materials were then extracted from the abstracts of these papers using Chemical Named Entity Recognition[1] to find a material that is rich in the discussion of impurity phases. $BiFeO_3$ was found to be the most data-rich, so we moved forward with these papers, 966 in total. Of this collection, 328 were determined to contain sol-gel processing procedures via a language-based processing paragraph classifier.[32] These were then manually inspected and verified based on the criteria that they contain: (1) thin film processing; (2) at least one target sample of $BiFeO_3$ without

21

dopants; (3) direct discussion of phase purity or impurity phase formation. This inspection yielded 121 articles. Additional articles were identified via a search over the Clarivate Analytics' Web of Science for articles related to BiFeO$_3$ fabrication from the years 2019-2021, since article volume from these years is limited in the text-mined datasets.[33] These were then inspected using the same criteria as above to yield the remaining 57 articles. Manual inspection and validation were performed by two human experts in text/data mining and materials science.

Processing procedure, outcome extraction and data cleaning: Manual extraction and validation of processing procedures and outcomes was preferred over automated methods due to the difficulty in accurately linking outcomes to respective procedures. All thin film processing procedures were extracted manually (from the same two experts mentioned above) into .csv format, where each column represents an aspect of the film fabrication (such as stirring temperature, annealing time, etc., as well as associated metadata) and each row represents a procedure and associated outcome. For uniformity of chemical and precursor naming, we ensured that each chemical name was standardized across its various representations (e.g. formulae were changed to the corresponding chemical name and alternative spellings were adjusted). All other parameters were numerical and standardized when necessary (e.g. descriptive times such as "one day" would be changed to "24 [hours]").

Data visualization: Visualizations to summarize the data in the text-mined dataset were created using the matplotlib (https://matplotlib.org/) and seaborn (https://seaborn.pydata.org/) Python visualization modules.

Film Fabrication: The precursor solution was prepared by dissolving Bi(NO$_3$)$_3$·5H$_2$O (≥99%, Sigma Aldrich) and Fe(NO$_3$)$_3$·9H$_2$O (≥99%, Sigma Aldrich) with 0.4 M concentration in 2-methoxyethanlol [2-ME] (anhydrous, 99.9%, Sigma-Aldrich), ethylene glycol [EG] (anhydrous, 99.9%, Sigma-Aldrich), or their mixture with volume ratio EG:2-ME (9:1) and (7:3). The

stoichiometric composition was varied from Bi: Fe = 0.9 to Bi: Fe = 1.1. For films prepared with chelating agents, either citric acid (99.9%, Sigma-Aldrich) or acetic acid (99.9%, Sigma-Aldrich) was added to precursor solution with molar ratio of Acid: Fe (4:1). After complete dissolution of the precursors, the solution was spin-coated on glass substrates at 3000 rpm for 30 sec, then dried on a hot plate at 80 °C for 10 min and baked on a hot plate at 300 °C for 10 min. The spin coating/baking procedure was repeated 5 times to obtain thick films. The as-cast baked films were annealed during in situ XRD measurements.

In situ X-ray diffraction (XRD): XRD measurements were performed at wavelength (1.5406 Å) from copper K-α x-ray radiation using Rigaku Smartlab X-ray diffractometer equipped with a HyPix-3000 high-energy-resolution multidimensional semiconductor detector, the temperature was increased from RT to 700 °C with an increment of 25 °C and was hold at each temperature for 5 min before doing XRD measurements (the heating/annealing profile is shown in the SI Fig. S7). Taking in situ diffraction measurements on a lab-based tool comes with the compromise of a small 2θ range versus potentially missing short-lived intermediate phases if scanning a longer 2θ range. Therefore, the in-situ measurements were conducted from 2θ = 10° - 35° with a scan rate of 3°/min and a step size of 0.1° (i.e. 8.3 min per pattern) to minimize the time for capturing the XRD pattern.

Computed thermodynamics: Standard Gibbs formation energies, $\Delta G_f(T)$, for CO_2 was obtained from NIST.[34] For solid-state compounds (i.e., $BiFeO_3$ and competing phases), formation enthalpies (at 0 K) were obtained with density functional theory (DFT) using the SCAN meta-GGA density functional.[25] Each initial structure was obtained from Materials Project[35] or the Inorganic Crystal Structure Database[36] and optimized using the Vienna *Ab Initio* Simulation Package (VASP)[37] and the projector augmented wave (PAW) method.[38] A plane wave energy cutoff of 520 eV and 1000 k-points per reciprocal atom were used for all calculations.

23

$\Delta G_f(T)$ for each solid-state compound was then obtained by combining the DFT-calculated formation enthalpies with the machine-learned descriptor introduced in ref.[26] $\Delta G_f(T)$ for CO_2 was obtained from NIST.[34] Gibbs formation energies were then used to compute all Gibbs reaction energies presented in this work, where the reaction energy is the difference between the sum of $\Delta G_f(T)$ for the products minus the sum of $\Delta G_f(T)$ for the reactants, weighted by the stoichiometric coefficients in the reaction. These reaction coefficients were determined on the basis of forming 1 mole of product (either $BiFeO_3$ or $Bi_2Fe_4O_9$). To compare the reaction energies for forming either product on even footing, the molar reaction energies were then normalized per atom in the product.

5. Acknowledgements

6. Author contributions

Conceptualization: M.A., K.H., K.C., G.C. and C.M.S.-F. Experimentation and analysis: M.A., K.H., and C.M.S.-F. Text mining: K.C., V.B., A.J., and G.C. First-principle calculations: C.J.B., A.T. and G.C. Supervision: G.C., A.J.,

C.M.S.-F. Writing—original draft: M.A., K.H. and C.M.S.-F. Writing—review and editing: K.C., C.J.B., A.T., V.B., A.J., G.C., and C.M.S.-F.

M.A. and K.H. contributed equally to this work

Competing interests: The authors declare that they have no competing interests.

7. References:

1   Hold, J. (2012). Materials genome initiative for global competitiveness (*National Science and Technology Council*), pp. 1-18.
2   Correa-Baena, J.-P., Hippalgaonkar,K., van Duren, J., Jaffer, S., Chandrasekhar, V.R., Stevanovic, V., Wadia, C., Guha, S., and Buonassisi, T. (2018). Accelerating materials development via automation, machine learning, and high-performance computing. Joule 2,1410–1420. https://doi.org/10.1016/j.joule.2018.05.009.
3   Eerenstein, W., Mathur, N.D. & Scott, J.F. (2006). Multiferroic and magnetoelectric materials. Nature 442, 759-765. https://doi.org/10.1038/nature05023.
4   Wang, J. . Neaton, B., Zheng, H., Nagarajan, V., Ogale, S.B., Liu, B., Viehland, D., Vaithyanathan, V., Schlom, D.G., Waghmare, U.V., et al. (2003). Epitaxial $BiFeO_3$ multiferroic thin film heterostructures. Science 299, 1719-1722. https://doi.org/10.1126/science.1080615.
5   Catalan, G., and Scott, J.F. (2009). Physics and applications of bismuth ferrite. Adv. Mater. 21, 2463-2485. https://doi.org/10.1002/adma.200802849.
6   Lebeugle, D., Colson, D., Forget, A., Viret, M., Bonville, P., Marucco, J.F., and Fusil, S. (2007). Room-temperature coexistence of large electric polarization and magnetic order in $BiFeO_3$ single crystals. *Phys. Rev. B* 76, 024116. https://doi.org/10.1103/Phys RevB.76.024116.
7   Anthony Raj, C., Muneeswaran, M., Jegatheesan, P., Giridharan, N.V., Sivakumar, V., and Senguttuvan, G. (2013). Effect of annealing time in the low-temperature growth of BFO thin films spin coated on glass substrates. J. Mater. Sci.: Mater. Electron. 24, 4148-4154. https://doi.org/10.1007/s10854-013-1374-0.
8   Sun, T., Hu, H., Pan, Z., Li, X., Wang, J., and Dravid, V.P. (2008). In situ real-time investigation of kinetics of nucleation and growth of sol-gel-derived functional oxide thin films. Phys. Rev. B 77, 205414. https://doi.org/10.1103/PhysRevB.77.205414.
9   Kononova, O., He, T., Huo, H., Trewartha, A., Olivetti, E.A., Ceder, G. (2021). Opportunities and challenges of text mining in materials research. iScience 24, 102155. https://doi.org/10.1016/j.isci.2021.102155.
10   Yang, B., Jin, L., Wei, R., Tang, X., Hu, L., Tong, P., Yang, J., Song, W., Dai, J., Zhu, X., et al. (2021). Chemical solution route for high-quality multiferroic $BiFeO_3$ thin films. Small 17, 1903663. https://doi.org/10.1002/smll.201903663.

11   Jo, S.-H., Lee, S.-G., and Lee, S.-H. (2012). Structural and pyroelectric properties of sol–gel derived multiferroic BFO thin films. Mater. Res. Bull. 47, 409-412. https://doi.org/10.1016/j.materresbull.2011.11.009.

12   Qi, X., Tsai, P.-C., Chen, Y.-C., Ko, C.-H., Huang, J.-C.-A., and Chen, I.-G. (2008). Ferroelectric properties and dielectric responses of multiferroic $BiFeO_3$ films grown by RF magnetron sputtering. J. Phys. D: App. Phys. 41, 232001. https://doi.org/10.1088/0022-3727/41/23/232001.

13   Wu, J., Wang, J., Xiao, D., and Zhu, J. (2011). Migration kinetics of oxygen vacancies in Mn-modified $BiFeO_3$ thin films. ACS Appl. Mater. Interfaces 3, 2504-2511. https://doi.org/10.1021/am2003747.

14   You, L., Chua, N. T., Yao, K., Chen, L., and Wang, J. (2009). Influence of oxygen pressure on the ferroelectric properties of epitaxial $BiFeO_3$ thin films by pulsed laser deposition. (2009). Phys. Rev. B 80, 024105. https://doi.org/10.1103/PhysRevB.80.024105.

15   Chen, P., Podraza, N.J., Xu, X.S., Melville, A., Vlahos, E., Gopalan, V., Ramesh, R., Schlom, D.G., Musfeldt, J.L., (2010). Optical properties of quasi-tetragonal $BiFeO_3$ thin films. Appl. Phys. Lett. 96, 131907. https://doi.org/10.1063/1.3364133.

16   Tang, X., Zhu, X., Dai, J., and Sun, Y. Self-limited grain growth, dielectric, leakage and ferroelectric properties of nanocrystalline $BiFeO_3$ thin films by chemical solution deposition. (2013). Acta Mater. 61, 1739-1747. https://doi.org/10.1016/j.actamat.2012.11.048.

17   Kartavtseva, M.S., Gorbenkoa, O.Y., Kaula, A.R., Murzinac, T.V., Savinovc, S.A., Barthélémy, A., (2007). $BiFeO_3$ thin films prepared using metalorganic chemical vapor deposition. Thin Solid Films 515, 6416-6421. https://doi.org/10.1016/j.tsf.2006.11.133.

18   Liu, H., and Wang, X. (2008). Large electric polarization in $BiFeO_3$ film prepared via a simple sol–gel process. J. Sol-Gel Sci. Technol. 47, 154-157. https://doi.org/10.1007/s10971-008-1773-5.

19   Guo, B. Deng, H., Zhai, X., Zhou, W., Meng, X., Weng, G., Chen, S., Yang, P., and Chu, J., (2017). Cr doping-induced structural phase transition, optical tuning and magnetic enhancement in $BiFeO_3$ thin films. Mater. Lett. 186, 198-201. https://doi.org/10.1016/j.matlet.2016.09.094.

20   Simões, A.Z., Cavalcante, L.S., Moura, F., Batista, N.C., Longo, E., and Varela J.A. (2012). Structure, microstructure, ferroelectric/electromechanical properties and retention characteristics of $[Bi_{1-x}Nb_x]FeO_3$ thin films. Appl. Phys. A 109, 703-714. https://doi.org/10.1007/s00339-012-7104-2.

21   Qi, X. Dho, J., Blamire, M., Jia, Q., Leem J.-S., Foltyn, S., MacManus-Driscoll, J.L. (2004). Epitaxial growth of $BiFeO_3$ thin films by LPE and sol–gel methods. J. Magn. Magn. Mater. 283, 415-421. https://doi.org/10.1016/j.jmmm.2004.06.014.

22   Liang, X.-L., Dai, J.-Q., Zhang, C.-C., and Cao, T.-F. (2021). Effects of solvents and Al doping on structure and physical properties of $BiFeO_3$ thin films. J. Sol-Gel Sci. Technol. 98, 45-53. https://doi.org/10.1007/s10971-021-05489-y.

23   Braunstein, G., Paz-Pujalt, G.R., Mason, M.G., Blanton, T., Barnes, C.L., and Margevich, D. (1993). The processes of formation and epitaxial alignment of $SrTiO_3$ thin films prepared by metallo-organic decomposition. *J. Appl. Phys.* 73, 961-970. https://doi.org/10.1063/1.353310.

24   Cho, S.G., Johnson, P.F., and Condrate, R.A. (1990). Thermal decomposition of (Sr, Ti) organic precursors during the Pechini process. J. Mater. Sci. 25, 4738-4744. https://doi.org/10.1007/BF01129934.

26

25 Sun, J., Ruzsinszky, A., and Perdew, J.P. (2015). Strongly constrained and appropriately normed semilocal density functional. Phys. Rev. Lett. 115, 036402. https://doi.org/10.1103/PhysRevLett.115.036402.

26 Bartel, C.J., Millican, S.L., Deml, A.M., Rumptz, J.R., Tumas, W., Weimer, A.W., Lany, S., Stevanović, V., Musgrave, C.B., and Holder, A.M. (2018). Physical descriptor for the Gibbs energy of inorganic crystalline solids and temperature-dependent materials chemistry. Nat. Commun. 9, 4168. https://doi.org/10.1038/s41467-018-06682-4.

27 Miura, A. Ito, H., Bartel, C.J., Sun, W., Rosero-Navarro, N.C., Tadanaga, K., Nakata, H., Maeda, K., and Cedercd, G. (2020). Selective metathesis synthesis of $MgCr_2S_4$ by control of thermodynamic driving forces. Mater. Horiz. 7, 1310-1316. https://doi.org/10.1039/C9MH01999E.

28 Bartel, C.J. (2022). Review of computational approaches to predict the thermodynamic stability of inorganic solids. J. Mater. Sci. 57, 10475-10498. https://doi.org/10.1007/s10853-022-06915-4.

29 Navarro, M.C., Lagarrigue, M.C., De Paoli, J.M., Carbonio, R.E., and Gómez, M.I. (2010). A new method of synthesis of $BiFeO_3$ prepared by thermal decomposition of $Bi[Fe(CN)_6]\cdot4H_2O$. J. Therm. Anal. Calorim. 102, 655-660. https://doi.org/10.1007/s10973-009-0417-5.

30 Miura, A., Bartel, C.J., Goto, Y., Mizuguchi, Y., Moriyoshi, C., Kuroiwa, Y., Wang, Y., Yaguchi, T., Shirai, M., Nagao, M., et al. (2021). Observing and modeling the sequential pairwise reactions that drive solid-state ceramic synthesis. Adv. Mater. 33, 2100312. https://doi.org/10.1002/adma.202100312.

31 Kononova, O., Huo, H., He, T., Rong, Z., Botari, T., Sun, W., Tshitoyan, V., and Ceder, G. (2019). Text-mined dataset of inorganic materials synthesis recipes. Sci. Data 6, 203. https://doi.org/10.1038/s41597-019-0224-1

32 He, T., Sun, W., Huo, H., Kononova, O., Rong, Z., Tshitoyan, V., Botari, T., Ceder, G., (2020). Similarity of Precursors in Solid-State Synthesis as Text-Mined from Scientific Literature. Chem. Mater. 32, 7861-7873. https://doi.org/10.1021/acs.chemmater.0c02553

33 Huo, H., Rong, Z., Kononova, O., Sun, W., Botari, T., He, T., Tshitoyan, V., and Ceder, G. (2019). Semi-supervised machine-learning classification of materials synthesis procedures. NPJ Comput. Mater. 5, 62. https://doi.org/10.1038/s41524-019-0204-1.

34 Chase, M.W., (1998). NIST-JANAF thermochemical tables. Vol. 9.

35 Jain, A., Ong, S.P., Hautier, G., Chen, W., Richards, W.D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., et al. (2013). ry: The Materials Project: A materials genome approach to accelerating materials innovation. APL Mater. 1, 011002. https://doi.org/10.1063/1.4812323

36 Belsky, A., Hellenbrandt, M., Karen, V., L.uksch, P. (2002). New developments in the Inorganic Crystal Structure Database (ICSD): accessibility in support of materials research and design. Acta Crystallogr. B: Struct. Sci. 58, 364-369. https://doi.org/10.1107/S0108768102006948.

37 Kresse, G., and Hafner, J. (1993). Ab initio molecular dynamics for liquid metals. (1993). Phys. Rev. B 47, 558-561. https://doi.org/10.1103/PhysRevB.47.558.

38 Kresse, G., and Joubert, D. (1999). From ultrasoft pseudopotentials to the projector augmented-wave method. Phys. Rev. B 59, 1758-1775. https://doi.org/10.1103/PhysRevB.59.1758