A Synthesizability-Guided Pipeline for Materials Discovery

Thorben Prein[§], **Willis O'Leary**[◇], **Aikaterini Flessa Savvidou**[◇], **Elchaïma Bourneix**[◇]
Joonatan E. M. Laulainen[◇]

[◇] Altrove, 25 Rue de Ponthieu, Paris, France
[§] Technische Universität München, Garching b. München, Germany

Abstract

Computational materials discovery relies on the generation of plausible crystal structures. The plausibility is typically judged through density functional theory methods which, while typically accurate at zero Kelvin, often favor low-energy structures that are not experimentally accessible. We develop a combined compositional and structural synthesizability score which provides an accurate way of predicting which compounds can actually be synthesized in a laboratory. We use it to evaluate non-synthesized structures from the Materials Project, GNoME, and Alexandria, and identified several hundred highly synthesizable candidates. We then predict synthesis pathways, conduct corresponding experiments, and characterize the products across 16 targets, **successfully synthesizing 7 of 16**. The entire experimental process was completed in only three days. Our results highlight omissions in lists of known synthesized structures, deliver insights into the practical utility of current materials databases, and showcase the central role synthesizability prediction can play in materials discovery.

1 Introduction

Discovery of new inorganic materials is a central goal of solid-state chemistry and, when achieved, can usher in enormous scientific and technological advancements. The development of machine-learning–accelerated, ultra-fast *in-silico* screening [Batzner et al., 2022, Chen and Ong, 2022, Wood et al., 2025, Rhodes et al., 2025] has unlocked vast databases of predicted candidate structures. Efforts such as the Materials Project [Horton et al., 2025], GNoME [Merchant et al., 2023], and Alexandria [Schmidt et al., 2023] contain structures predicted using active learning and a mix of composition- and structure-based models to propose candidate structures, which are subsequently assessed by density functional theory (DFT) simulations. Analogous to virtual libraries in drug discovery [Zdrazil et al., 2024, Tingle et al., 2023], these resources provide initial property estimates alongside full crystallographic descriptions of putative phases.

The number of proposed inorganic crystals now exceeds the number of experimentally synthesized compounds by more than an order of magnitude [Belsky et al., 2002]. The current challenge is to determine which of these predicted materials can be fabricated. This has historically been addressed by computing convex-hull stabilities with DFT [Lee et al., 2022]. While this approach constitutes a useful first filter, it typically overlooks finite-temperature effects, namely entropic and kinetic factors, that govern synthetic accessibility [Bartel, 2022]. As a result, it has become increasingly difficult to distinguish purported stable structures from truly synthesizable ones. For example, the Materials Project lists 21 SiO_2 structures within 0.01 eV of the convex hull (as of August 2025). The second most common phase SiO_2, cristobalite (β-quartz), is not among these 21. Overall, there is a pressing need for accurate synthesizability assessments to efficiently steer scientists and engineers toward compounds that are readily accessible in the laboratory [Park et al., 2025].

39th Conference on Neural Information Processing Systems (NeurIPS 2025) AI4Mat Workshop.

Given the abundance of predicted near-stable crystal structures, procedures are needed to decide which candidates to attempt to synthesize and under which conditions these syntheses should be carried out (process parameters). So far, several data-driven approaches have been introduced, including graph-based structural models and composition-based models [Lee et al., 2022, Antoniuk et al., 2023, Song et al., 2025, Davariashtiyani et al., 2021, Zhu et al., 2023, Gleaves et al., 2023, Amariamir et al., 2025]. These approaches consider composition and structure in isolation and, to date, have not convincingly demonstrated that they can drive the experimental synthesis of novel compounds. Here, we propose and implement a unified, synthesis-aware prioritization framework that integrates complementary signals from composition and crystal structure. Composition signals may be governed by elemental chemistry, precursor availability, redox and volatility constraints, whereas structural signals capture local coordination, motif stability, and packing. Our unified model demonstrates state-of-the-art performance. We apply our framework to more than 4.4 million simulated crystal structures, yielding 24 candidates predicted to be highly synthesizable. Subsequently, we employ neural models trained on literature-mined synthesis recipes to predict process parameters and subsequently carry out syntheses on the proposed candidates in an automated solid-state laboratory. Of the 16 samples we successfully characterized, seven matched the target structure, including one completely novel and one previously unreported structure.

2 Methods

Fig. 1 outlines our synthesizability-guided pipeline. We first screened a pool of 4.4 million computational structures, of which 1.3 million were calculated to be synthesizable (detailed distributions Fig, 4 and Fig. 5). We then focused only on highly synthesizable (0.95 rank-average per our synthesizability score) materials and removed compounds containing platinoid group elements, yielding approximately 15,000 candidates. Finally, we reduced the candidates by removing non-oxides and toxic compounds, yielding approximately 500 final structures. We applied retrosynthetic planning on these prioritized structures to generate feasible routes, of which a subset (see Section 2.2) was then executed in a high-throughput laboratory platform. The resulting products were verified automatically by X-ray diffraction (XRD) (see Appendix for details).

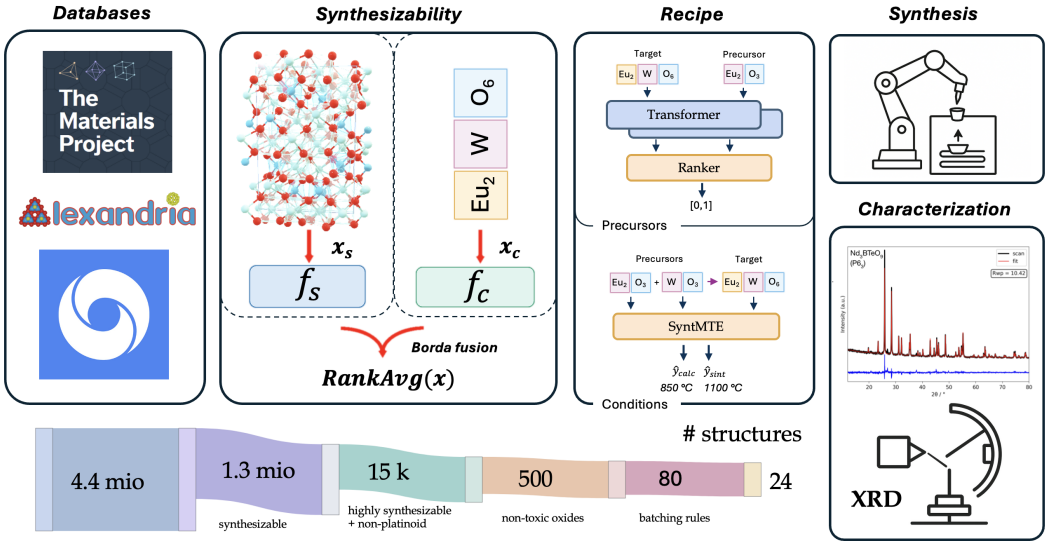

Figure 1: Our pipeline for selecting synthesizable candidates, predicting their recipe, carrying out the synthesis, and characterizing their structure.

2.1 Synthesizability Model

We define synthesizability as the probability that a compound can be prepared in the laboratory using currently available synthetic methods. Machine-learning efforts to predict this outcome largely fall into two families: (i) composition-only models that operate on stoichiometry or engineered

composition descriptors [Lee et al., 2022, Antoniuk et al., 2023]; and (ii) structure-aware models that leverage crystal-structure graphs [Song et al., 2025, Davariashtiyani et al., 2021, Zhu et al., 2023, Gleaves et al., 2023, Amariamir et al., 2025].

Problem Formulation Let each candidate material be represented by its composition x_c and a relaxed crystal structure x_s. We denote the input as $x = (x_c, x_s)$ and the label $y \in \{0, 1\}$ indicating whether the compound has been experimentally synthesized. The goal is to learn a score $s(x) \in [0, 1]$ that estimates a compound's synthesizability.

Data Curation We built our training dataset from the Materials Project [Jain et al., 2013]. This ensures consistency between composition and structure in order to obtain supervision without inheriting artifacts common in experimental entries encountered e.g. in the Inorganic Crystal Structure Database (ICSD) [Hellenbrandt, 2004], such as non-stoichiometry, dopants or partial occupancies. We assigned labels according to the Materials Project's "theoretical" field, a flag which corresponds to whether ICSD entries exist for a given structure. We labeled a composition as unsynthesizable ($y = 0$) if all polymorphs of that composition are flagged as theoretical. Conversely, we labeled a composition as synthesizable ($y = 1$) if there exists any polymorph that is not flagged as theoretical. The final dataset comprised 49,318 synthesizable compositions and 129,306 unsynthesizable compositions, stratified into train/validation/test splits.

Model Our model integrates complementary signals from composition and structure via two encoders, learning a binary classification task:

$$\mathbf{z}_c = f_c(x_c;\, \theta_c), \qquad \mathbf{z}_s = f_s(x_s;\, \theta_s),$$

where f_c is a fine-tuned compositional MTEncoder transformer [Prein et al., 2023] and f_s is a graph neural network fine-tuned from the JMP model [Shoghi et al., 2023] on the crystal structure x_s. Both encoders are extensively pretrained and each feeds a small MLP head that outputs a separate synthesizability score. During training, we fine-tune all parameters end-to-end on an NVIDIA H200 cluster. We minimize binary cross-entropy with early stopping on validation AUPRC.

Usage in Screening At inference, each model outputs a synthesizability probability for candidate x_i. To obtain an enhanced ranking across candidates, we aggregate both predictions via a rank-average ensemble (Borda fusion). Specifically, we convert probabilities to ranks and define

$$\mathrm{RankAvg}(i) = \frac{1}{2N} \sum_{m \in \{c,s\}} \left(1 + \sum_{j=1}^{N} \mathbf{1}\big[s_m(j) < s_m(i)\big] \right).$$

Here $i \in \{1, \ldots, N\}$ indexes candidates. N is the total number of candidates in the screening pool. The index $m \in \{c, s\}$ refers to the composition and structure models, and $s_m(i) \in [0, 1]$ is the synthesizability probability predicted by model m for candidate i. By construction, $\mathrm{RankAvg}(i) \in [1/N, 1]$, with larger values indicating greater predicted synthesizability. Consequently, we rank candidates by $\mathrm{RankAvg}$ rather than applying probability thresholds.

Synthesis Planning After identifying high-priority candidates with the rank-based screening, we generated synthesis recipes in two stages. First, we applied Retro-Rank-In [Prein et al., 2025a], a precursor-suggestion model, to produce a ranked list of viable solid-state precursors for each target. Second, we selected the top-ranked precursor pairs and use SyntMTE [Prein et al., 2025b] to predict the calcination temperature required to form the target phase. We then balanced the reaction and compute the corresponding precursor quantities. Both models are trained on literature-mined corpora of solid-state synthesis [Kononova et al., 2019].

2.2 Experimental Synthesis and Characterization

Of the ca. 500 candidates shown in Fig. 1, we selected the experimental candidates by a combination of a web-searching LLM to judge whether each candidate was likely to have been synthesized before, followed by an expert judgment to remove targets with unrealistic oxidation states given an oxygenated furnace (e.g. $BaTe_2O$) and those with common formulas (e.g. DyO_2) as these compounds are likely to be well-explored. This yielded a final list of 80 targets.

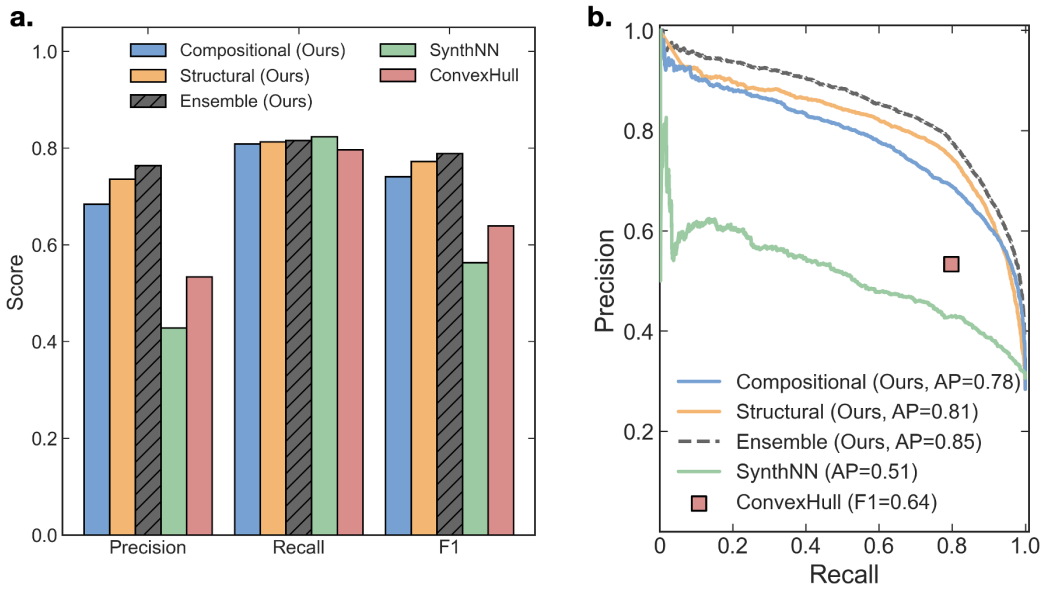

Figure 2: Test-set performance of the compositional, structural, and ensemble models. **(a)** Precision, recall, and F1 scores for identifying the target "theoretical" labels in the materials project. **(b)** ROC curves (AUC shown in the legend). The convex-hull baseline classifies compounds within $50\,\mathrm{meV}$ above the hull as synthesizable.

Next, based on recipe similarity, we automatically selected 24 targets across two batches of 12, such that all 12 experiments in one batch could be carried out in the furnace at the same time. Finally, the samples were weighed, ground and calcined in a Thermo Scientific Thermolyne Benchtop Muffle Furnace in our high-throughput laboratory (see Appendix for further experimental details). Eight of the 24 samples strongly bonded to their crucibles during synthesis and therefore could not be subjected to further characterization. The remaining 16 samples could be milled after synthesis and were subsequently characterized using a Malvern Panalytical Aeris benchtop X-ray diffractometer. The X-ray spectra were automatically analyzed, then expert analysis was used to confirm the presence of the target compared to existing phases.

3 Results

3.1 Model Performance

We benchmark our approach against two baselines: SynthNN, a recent structure-aware synthesizability model Antoniuk et al. [2023], and a convex-hull classifier that labels compounds as synthesizable when their Materials Project energy above the hull is $\leq 50\,\mathrm{meV/atom}$. While both baselines attain competitive recall, they suffer from markedly lower precision than our method, which causes them to over-prioritize many compositions that are unlikely to be synthesizable (Fig. 2). Specifically, SynthNN achieves $R = 0.823$ but only $P = 0.428$, and the convex-hull heuristic reaches $R = 0.797$ with $P = 0.534$.

To isolate the contributions inside our ensemble approach, we compare the compositional and the structural part with the final rank–average ensemble. As summarized in Fig. 2 (a,b), the structural model improves upon the compositional baseline across all metrics $P = 0.736$ vs. 0.684, $R = 0.813$ vs. 0.809, and $F_1 = 0.773$ vs. 0.741 and the ensemble performs best overall, reaching $P = 0.764$, $R = 0.816$, and $F_1 = 0.789$. Recall remains tightly clustered across our three variants, indicating that the bulk of improvement comes from higher precision. Since our final ensemble ranking lacks a meaningful global threshold, we calibrate each model's threshold to maximize its F1 score to ensure comparability.

We hypothesize that the compositional model's competitiveness reflects that much of the synthesizability signal resides in elemental chemistry rather than exact atomic geometry, charge balance, valence matching, redox windows, and volatility constraints are inherently compositional. These signals recover many true positives from the unlabeled pool, boosting recall, but also inflate false positives for stoichiometries that appear feasible yet cannot form stable coordination networks. By encoding local coordination and connectivity, the structural model suppresses precisely these spurious cases, recovering precision with minimal cost in recall. The final rank–average ensemble inherits the compositional coverage while recapturing most of the lost precision, yielding the best overall balance.

Case Illustrations Let s_{comp} denote the compositional prediction and s_{struct} the structural prediction for the *same* candidate structure.

- *N_2*: $s_{\text{comp}} = 0.709$, $s_{\text{struct}} = 1.10 \times 10^{-4}$. While the composition looks plausible in isolation, with nitrogen-rich formulas occurring near many synthesized nitrides, the candidate is a molecular solid likely requiring non-ambient conditions. The structure encoder down-weights such graphs because they lack robust inorganic connectivity under our solid-state synthesis domain.
- *$Cu_6Te_2Mo_2H_2Cl_4O_{15}$*: $s_{\text{comp}} = 0.858$, $s_{\text{struct}} = 1.66 \times 10^{-3}$. The composition sits near known oxyhalides and oxychalcogenides, leading to high s_{comp}, but the proposed structure mixes competing anions and volatile species that are difficult to stabilize in a single phase via high-temperature calcination. The structure model captures this through unfavorable coordination and packing cues, yielding a low s_{struct}.

Because our supervision stems from positive-unlabeled (PU) entries with reports of synthesized materials being treated as positives, while all others are unlabeled rather than true negatives, the precision values above should be viewed as conservative. In PU settings, many predicted positives inside the unlabeled pool are plausible true positives, thus the observed increases in precision from structure and ensembling represent lower bounds on their practical utility. Conversely, recall is computed on labeled positives and is therefore less sensitive to unlabeled contamination, explaining the narrow spread in recall across models.

3.2 Experimental Results

The targets, synthesis results, along with an example XRD spectrum and structure, are provided in Fig. 3. Of the 16 characterized samples, we successfully synthesized the target seven times, (44%), synthesized a predicted or known polymorphs two times (12%), and failed to synthesize the target 7 times (44%) (Fig. 3a). Of the nine syntheses where either the target or a polymorph was synthesized, three were subsequently identified to be nearly identical to a structure already in the Materials Project (and ICSD) that is listed as synthesized. These structures were therefore effectively present in the training data. Of the remaining six, a similar structure was present in the ICSD for three compounds, although these compounds were erroneously classed as theoretical in the Materials Project database.

We synthesized three compounds with no known structure reported in ICSD:

- Y_2MnFeO_6 (Alexandria): Three polymorphs with the composition have been synthesized, while one theoretical structure is available both in Materials Project and Alexandria. We detected our target polymorph at a low (3.3%) concentration in addition to observing a phase mixture of the other two known polymorphs (see Appendix for XRD fit).
- Nd_3BTeO_9 (GNoME): We found a paper listing its synthesis but without a CIF file [Zhou and Li, 2023]. Our refined structure is isostructural to Bi_3BTeO_9, with a $P6_3$ space group and $a = 8.761$ Å and $c = 5.553$ Å. We achieved near full purity, with $< 5\%$ of intensity accounted for by an unknown phase that we could not identify.
- Eu_2WO_6 (Alexandria): A CIF file for this structure exists in the ICSD, but with significantly (ca. 5%) different lattice parameters. The observed orthorhombic polymorph was different than the original target and was formed with ca. 88% purity.

a.

Target phase	Target formed	Novel	Comments
Y_2MnFeO_6 [A]	✅	✅	3% purity
Nd_3BTeO_9 [G]	✅	~	Reported but no CIF
Eu_2WO_6 [A]	~	~	Formed polymorph
$DyAl_3(BO_3)_4$ [A]	✅	✗*	60% purity
$Nd_3Mn_2Sb_3O_{14}$ [A]	✅	✗*	
Y_2MnCoO_6 [A]	✅	✗*	
Sm_3BWO_9 [A]	✅	✗*	
$Sr_4Al_6MoO_{16}$ [G]	~	✗*	Formed polymorph
$TbFeO_3$ [A]	✅	✗*	40 % purity
Ba_2DyFeO_5 [G]	✗	✅	
$Ba_2Gd(CuO_2)_4$ [M]	✗	✅	
Ba_2TmFeO_5 [G]	✗	✅	
Eu_3MoO_7 [A]	✗	✅	
$La_2Te_2WO_{10}$ [G]	✗	✅	
Nd_3BMoO_9 [G]	✗	✅	
$Ho_2Mn_2O_7$ [M]	✗	✗*	

[M]Materials Project | [G]GNoME | [A]Alexandria | *reported in ICSD

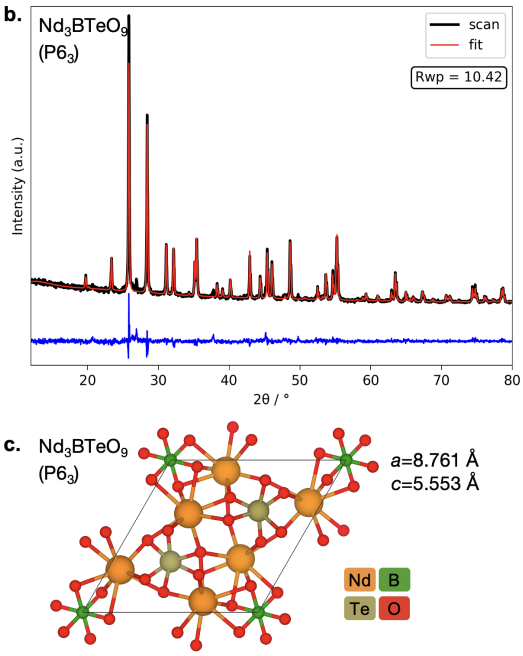

Figure 3: **a.** Identified highly synthesizable ($\mathrm{RankAvg} > 0.95$) targets and the results of performed experiments. For Eu_2WO_6, a non-target polymorph was synthesized that had the same space group as a known structure but with significantly different lattice parameters. For $Sr_4Al_6MoO_{16}$, a compound with the target composition was synthesized. However, due to low purity we could not rule out formation of one of the three known polymorphs. Several compounds, while listed as theoretical in Materials Project, do have similar structures reported, mostly in the ICSD. **b.** XRD scan, fit, and difference for Nd_3BTeO_9. A second unknown phase is present and accounts for <5% of total intensity. **c.** Refined hexagonal unit cell for Nd_3BTeO_9. The unit cell is hexagonal with the lattice parameters $a = 8.761$ Å and $c = 5.553$ Å

4 Discussion

4.1 Synthesis of Novel Compounds

The successful one-shot synthesis of (a) novel compounds and (b) known compounds excluded from our training set highlights the promise of our approach as well as the broader value of machine-learning-based synthesizability screening methods. Given current trends, we expect the number and diversity of predicted inorganic structures to continue to grow. We expect many of these structures to be stable according to DFT-derived metrics but ultimately unsynthesizable due to high kinetic barriers or stability only under unusual or wholly impractical conditions. With our approach, we enjoyed a 44% success rate in selecting synthesizable structures from millions of possible candidates. Presumably, some candidates in the remaining 56% may be synthesizable under different conditions or through the use of alternative synthesis methods.

Still, our successes carry with them strong caveats. First, most of the successful synthesis candidates were isostructural to other known compounds. This might reflect our greedy selection of candidates with only the highest synthesizability probabilities. It may also arise from data limitations and the challenges to build models that generalize to exotic, never-before-seen structures. Furthermore, many of our successful syntheses were in fact not novel compounds, despite being marked as purely theoretical in Materials Project. Clearly, our automated screening to find unsynthesized compounds can be significantly improved. Although our results are encouraging as far as model performance is concerned, it also motivates efforts to provide more comprehensive and accurate data for training synthesizability models.

Finally, we would like to emphasize the importance of both the structural and compositional components of the model. In preliminary work, we attempted the synthesis of nine candidates selected using a composition-only model but formed none of the targets. Upon evaluating the candidates, we noted that the majority were unrealistic polymorphs of well-established compounds. In addition, we observed a tendency to up-rank materials with complex ordering of transition metals (e.g. $ScTa_5Nb_6Tc_{12}$). These structures are rarely experimentally achieved due to entropic effects but tend to be overrepresented in DFT-derived materials datasets. These findings motivated trials with a purely structure-based model. The use of a structure-based model eliminated the aforementioned candidates but favored compounds with unrealistic chemistries (e.g. AsH_4F_7). In the end, synthesizing the complementary signals of composition-based and structure-based models led to superior overall performance in predicting synthesizability, in this context, all of the candidates we tried in the lab achieved high scores in both modalities.

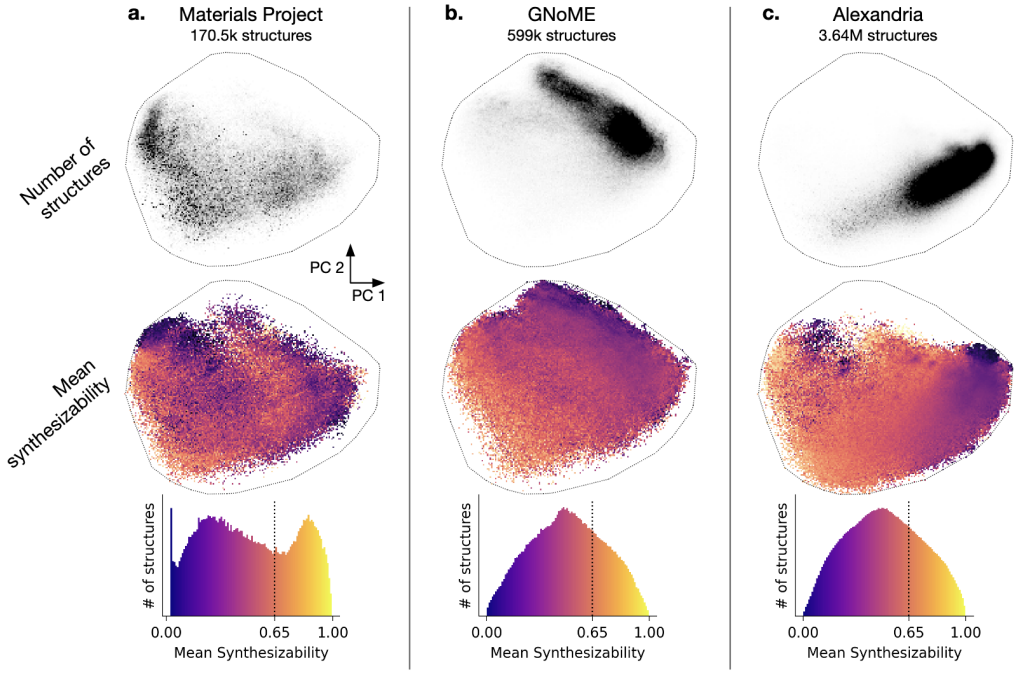

Figure 4: Joint principal component analyses on MTEncoder embeddings shown in the same latent space for **a.** Materials Project, **b.** GNoME, and **c.** Alexandria. Top row: density of structures by embedding. Middle row: average synthesizability by embedding. Bottom row: distribution of rank-average synthesizabilities within each dataset.

4.2 Dataset-level Insights

With a robust score for synthesizability in hand, we can better understand the distribution of synthesizabilities for the roughly 4.4 million materials available in Materials Project, GNoME, and Alexandria. To do this, we generated MTEncoder [Prein et al., 2023] embeddings for all compositions contained within these datasets then carried out principal component analysis (PCA). Dataset metrics are plotted with respect to the first two principal components in Fig. 4. We find that these datasets cover quite distinct regions in principal component space. Materials Project contains a relatively diverse set of materials, distributed across the space, while the materials contained in GNoME and Alexandria are concentrated in comparatively narrow regions.

For each database, we computed mean synthesizability across principal component space. Certain regions in Materials Project boast fairly high average synthesizability while others do not. The existence of "patches" of synthesizability is unsurprising, as we would expect a complex relationship between principal components and any material property. For GNoME and Alexandria, we observe a roughly inverse relationship between the number of materials and mean synthesizability in a given

region. A number of possible factors could give rise to this trend. One intuitive reason is that, in the drive to explore more diverse structures and compositions, the GNoME and Alexandria projects have generated less synthesizable materials. For example, compounds containing a greater number of elements are less likely to have been synthesized, on average. In addition, the simplest structures and compositions have likely already been synthesized, and therefore are not present in GNoME or Alexandria. Finally, this trend may simply reflect differences in the data distribution between Materials Project – the source of our positive labels – and other databases. Overall, this highlights the need to critically judge predicted crystal structures for synthesizability, especially as the total number of predicted structures continues to grow.

4.3 Data Limitations

As noted above, data limitations, specifically a lack of a fully unified and accurate database containing all known synthesized structures combining published literature, ICSD, and Materials Project entries result in our most synthesizable structures being already synthesized. Additional unification of this database, as well as separation of whether such structures have been made using solid-state or hydrothermal methods, and under ambient or high pressures would enable the development of highly accurate solid-state synthesizability score. This is especially important as solid-state methods make up the bulk of industrial processing for inorganic compounds. Accordingly, we treat the Materials Project "theoretical" designation as our proxy for a discovery setting.

5 Conclusion

Crystal generation could enable a paradigm shift in the development of novel materials, but judging the plausibility of structures is increasingly important as the number of generated plausible structures increases. Our synthesizability-guided pipeline for materials discovery was able to reliably identify both novel materials and recover materials mislabeled as theoretical in our datasets. Of the 16 targets, we successfully synthesized seven, formed two polymorphs, and failed in the remaining seven cases. Our high success rate (44%) was achieved with only one experiment per target.

To our knowledge, this work is among the first demonstrations in which machine-learned synthesizability predictions guided experimental target selection and resulted in previously unreported compounds. Starting from a corpus of 4.4 million candidate structures, our integrated screening prioritized candidates with high predicted synthesizability. Focusing effort on these targets minimized time spent on unrealistic compositions and accelerated laboratory progress toward highest priority materials. Together, these results underscore the practical impact of synthesizability-aware screening in directing material discovery.

References

Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature communications*, 13(1):2453, 2022.

Chi Chen and Shyue Ping Ong. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, 2022.

Brandon M Wood, Misko Dzamba, Xiang Fu, Meng Gao, Muhammed Shuaibi, Luis Barroso-Luque, Kareem Abdelmaqsoud, Vahe Gharakhanyan, John R Kitchin, Daniel S Levine, et al. Uma: A family of universal models for atoms. *arXiv preprint arXiv:2506.23971*, 2025.

Benjamin Rhodes, Sander Vandenhaute, Vaidotas Šimkus, James Gin, Jonathan Godwin, Tim Duignan, and Mark Neumann. Orb-v3: atomistic simulation at scale. *arXiv preprint arXiv:2504.06231*, 2025.

Matthew K Horton, Patrick Huck, Ruo Xi Yang, Jason M Munro, Shyam Dwaraknath, Alex M Ganose, Ryan S Kingsbury, Mingjian Wen, Jimmy X Shen, Tyler S Mathis, et al. Accelerated data-driven materials science with the materials project. *Nature Materials*, pages 1–11, 2025.

Amil Merchant, Simon Batzner, Samuel S. Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, 2023. doi: 10.1038/s41586-023-06735-9.

Jonathan Schmidt, Noah Hoffmann, Hai-Chen Wang, Pedro Borlido, Pedro JMA Carriço, Tiago FT Cerqueira, Silvana Botti, and Miguel AL Marques. Machine-learning-assisted determination of the global zero-temperature phase diagram of materials. *Advanced Materials*, 35(22):2210788, 2023.

Barbara Zdrazil, Eloy Felix, Fiona Hunter, Emma J Manners, James Blackshaw, Sybilla Corbett, Marleen De Veij, Harris Ioannidis, David Mendez Lopez, Juan F Mosquera, et al. The chembl database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic acids research*, 52 (D1):D1180–D1192, 2024.

Benjamin I Tingle, Khanh G Tang, Mar Castanon, John J Gutierrez, Munkhzul Khurelbaatar, Chinzorig Dandarchuluun, Yurii S Moroz, and John J Irwin. Zinc-22- a free multi-billion-scale database of tangible compounds for ligand discovery. *Journal of chemical information and modeling*, 63(4):1166–1176, 2023.

Alec Belsky, Mariette Hellenbrandt, Vicky Lynn Karen, and Peter Luksch. New developments in the inorganic crystal structure database (icsd): accessibility in support of materials research and design. *Structural Science*, 58(3):364–369, 2002.

Andrew Lee, Suchismita Sarker, James E Saal, Logan Ward, Christopher Borg, Apurva Mehta, and Christopher Wolverton. Machine learned synthesizability predictions aided by density functional theory. *Communications Materials*, 3(1):73, 2022.

Christopher J Bartel. Review of computational approaches to predict the thermodynamic stability of inorganic solids. *Journal of Materials Science*, 57(23):10475–10498, 2022.

Hyunsoo Park, Kinga O Mastej, Panyalak Detrattanawichai, Ryan Nduma, and Aron Walsh. Closing the synthesis gap in computational materials design. 2025.

Evan R Antoniuk, Gowoon Cheon, George Wang, Daniel Bernstein, William Cai, and Evan J Reed. Predicting the synthesizability of crystalline inorganic materials from the data of known material compositions. *npj Computational Materials*, 9(1):155, 2023.

Zhilong Song, Shuaihua Lu, Minggang Ju, Qionghua Zhou, and Jinlan Wang. Accurate prediction of synthesizability and precursors of 3d crystal structures via large language models. *Nature Communications*, 16(1): 6530, 2025.

Ali Davariashtiyani, Zahra Kadkhodaie, and Sara Kadkhodaei. Predicting synthesizability of crystalline materials via deep learning. *Communications Materials*, 2(1):115, 2021.

Ruiming Zhu, Siyu Isaac Parker Tian, Zekun Ren, Jiali Li, Tonio Buonassisi, and Kedar Hippalgaonkar. Predicting synthesizability using machine learning on databases of existing inorganic materials. *ACS omega*, 8(9):8210–8218, 2023.

Daniel Gleaves, Nihang Fu, Edirisuriya M Dilanga Siriwardane, Yong Zhao, and Jianjun Hu. Materials synthesizability and stability prediction using a semi-supervised teacher-student dual neural network. *Digital Discovery*, 2(2):377–391, 2023.

Sasan Amariamir, Janine George, and Philipp Benner. Syncotrain: a dual classifier pu-learning framework for synthesizability prediction. *Digital Discovery*, 4(6):1437–1448, 2025.

Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1), 2013.

Mariette Hellenbrandt. The inorganic crystal structure database (icsd)—present and future. *Crystallography Reviews*, 10(1):17–22, 2004.

Thorben Prein, Elton Pan, Tom Doerr, Elsa Olivetti, and Jennifer LM Rupp. Mtencoder: A multi-task pretrained transformer encoder for materials representation learning. In *AI for Accelerated Materials Design-NeurIPS 2023 Workshop*, 2023.

Nima Shoghi, Adeesh Kolluru, John R Kitchin, Zachary W Ulissi, C Lawrence Zitnick, and Brandon M Wood. From molecules to materials: Pre-training large generalizable models for atomic property prediction. *arXiv preprint arXiv:2310.16802*, 2023.

Thorben Prein, Elton Pan, Sami Haddouti, Marco Lorenz, Janik Jehkul, Tymoteusz Wilk, Cansu Moran, Menelaos Panagiotis Fotiadis, Artur P Toshev, Elsa Olivetti, et al. Retro-rank-in: A ranking-based approach for inorganic materials synthesis planning. *arXiv preprint arXiv:2502.04289*, 2025a.

Thorben Prein, Elton Pan, Janik Jehkul, Steffen Weinmann, Elsa A Olivetti, and Jennifer LM Rupp. Language models enable data-augmented synthesis planning for inorganic materials. *arXiv preprint arXiv:2506.12557*, 2025b.

Olga Kononova, Haoyan Huo, Tanjin He, Ziqin Rong, Tiago Botari, Wenhao Sun, Vahe Tshitoyan, and Gerbrand Ceder. Text-mined dataset of inorganic materials synthesis recipes. *Scientific data*, 6(1):203, 2019.

Changqing Zhou and Rukang Li. Large difference in nonlinear optical activity of rare earth ion substitution of bi3+ in a3tebo9 (a= bi, la, pr, nd, sm-dy). *Inorganic Chemistry*, 62(28):11265–11270, 2023.

Nicola Doebelin and Reinhard Kleeberg. Profex: a graphical user interface for the rietveld refinement program bgmn. *Applied Crystallography*, 48(5):1573–1580, 2015.

A Technical Appendices and Supplementary Material

Symbol	Meaning / Type
x_c	Composition (stoichiometric descriptor/vector over elements)
x_s	Relaxed crystal structure (atomic positions + lattice); input structure
$x = (x_c, x_s)$	Full input for a candidate material
$y \in \{0, 1\}$	Label: 1 if structure is reported synthesized; 0 otherwise
$f_c(\cdot; \theta_c)$	Composition encoder (MTEncoder); parameters θ_c
$f_s(\cdot; \theta_s)$	Structure encoder (GNN fine-tuned from JMP); parameters θ_s
$\mathbf{z}_c$	Composition embedding $= f_c(x_c; \theta_c) \in \mathbb{R}^{d_c}$
$\mathbf{z}_s$	Structure embedding $= f_s(x_s; \theta_s) \in \mathbb{R}^{d_s}$
$s_c(x)$	Synthesizability probability from the composition head, in $[0, 1]$
$s_s(x)$	Synthesizability probability from the structure head, in $[0, 1]$

Table 1: Notation used in the ML formulation and evaluation.

A.1 Classification Metrics

We cast synthesizability prediction as a class-imbalanced binary classification task with definitive positive and negative labels. Let $\mathrm{TP}, \mathrm{FP}, \mathrm{TN}, \mathrm{FN}$ denote true/false positives/negatives computed at a decision threshold τ on the predicted probability $s(x) \in [0, 1]$.

Thresholded metrics. We report precision, recall (a.k.a. true positive rate), and their harmonic mean F_1:

$$\text{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \qquad \text{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \qquad F_1 = \frac{2\,\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}.$$

A.2 Additional Synthesizability Prediction Results

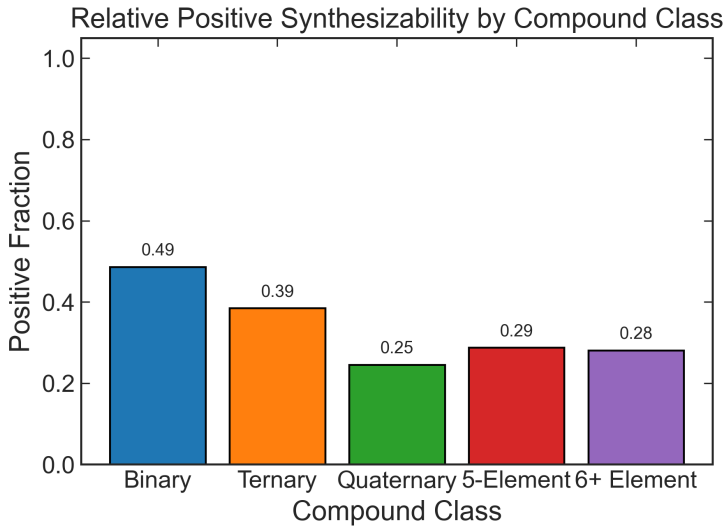

Figure 5: Synthesizability by compound class. Computed from entries in the Materials Project.

A.3 Additional Experimental Details

The experiments were conducted in our high-throughput laboratory. The proportions of precursor materials were selected to match stoichiometry without consideration for the differing volatility of different precursors. As some precursors are more volatile and therefore disproportionately more 'lost' during the reaction, some purity is likely lost in this non-optimized process. All precursors were combined together and milled for 10

[0]Here $s(i)$ denotes the model probability used in the loss (e.g., per-head); later ranking uses $\mathrm{RankAvg}$.

Target phase	Precursors	Temperature	Time
Y_2MnFeO_6 [A]	Y_2O_3, Mn_2O_3, Fe_3O_4	1030°C	8h
Nd_3BTeO_9 [G]	Nd_2O_3, B_2O_3, TeO_2	965°C	8h
Eu_2WO_6 [A]	Eu_2O_3, WO_3	1030°C	8h
$DyAl_3(BO_3)_4$ [A]	Dy_2O_3, Al_2O_3, B_2O_3	965°C	8h
$Nd_3Mn_2Sb_3O_{14}$ [A]	Nd_2O_3, MnO_2, Sb_2O_3	1030°C	8h
Y_2MnCoO_6 [A]	Y_2O_3, Mn_2O_3, Co_3O_4	1030°C	8h
Sm_3BWO_9 [A]	Sm_2O_3, B_2O_3, WO_3	965°C	8h
$Sr_4Al_6MoO_{16}$ [G]	$SrCO_3$, Al_2O_3, MoO_3	1030°C	8h
$TbFeO_3$ [A]	Tb_4O_7, Fe_2O_3	1030°C	8h
Ba_2DyFeO_5 [G]	$BaCO_3$, Dy_2O_3, Fe_2O_3	965°C	8h
$Ba_2Gd(CuO_2)_4$ [M]	$BaCO_3$, Gd_2O_3, CuO	1030°C	8h
Ba_2TmFeO_5 [G]	$BaCO_3$, Tm_2O_3, Fe_2O_3	965°C	8h
Eu_3MoO_7 [A]	Eu_2O_3, MoO_3	1030°C	8h
$La_2Te_2WO_{10}$ [G]	La_2O_3, TeO_2, WO_3	965°C	8h
Nd_3BMoO_9 [G]	Nd_2O_3, $B(OH)_3$, MoO_3	1030°C	8h
$Ho_2Mn_2O_7$ [M]	Ho_2O_3, MnO_2	1030°C	8h

Figure 6: Experimental details for the 16 targets, including precursors, processing temperatures, and times.

minutes using a Hauschild Speedmixer Smart Dac 400.3 FVZ to reduce particle size and thoroughly mix the components. Subsequently, the powders were placed in alumina crucibles and placed in a Thermo Scientific Thermolyne Benchtop Muffle Furnace for a specified time at a specified temperature. The precursors, times, and temperatures are listed below in Fig. 6. The samples were subsequently characterized using a Malvern Panalytical Aeris benchtop X-ray diffractometer. The fitting of the XRD spectra was done using an in-house code package that uses a BGMN Rietveld refinement as a final step Doebelin and Kleeberg [2015]. The exact fitting procedure is discussed in the Appendix.

Originally, there were 24 targets, of which 8 samples, while fully undergoing synthesis, were not able to be extracted out of their crucibles for XRD analysis. These samples will subsequently be processed and evaluated using the same pipeline, and as such, we do not consider them failed attempts, but rather attempts in progress.

A.4 XRD Fits for Experiments

To perform the XRD fits, we used an in-house phase-identification algorithm that picked likely phases from the ICSD, Materials Project, GNoME, and Alexandria. These were weighted based on thermodynamic likelihood. These were then human-validated, before an automated Rietveld refinement procedure was performed using BGMN [Doebelin and Kleeberg, 2015], a robust automated Rietveld refinement algorithm. These results were compared against alternative models with different phases to perform an additional qualitative assessment of the model fit.

We provide XRD fits for all of the claimed successful syntheses below.

Y_2MnFeO_6:

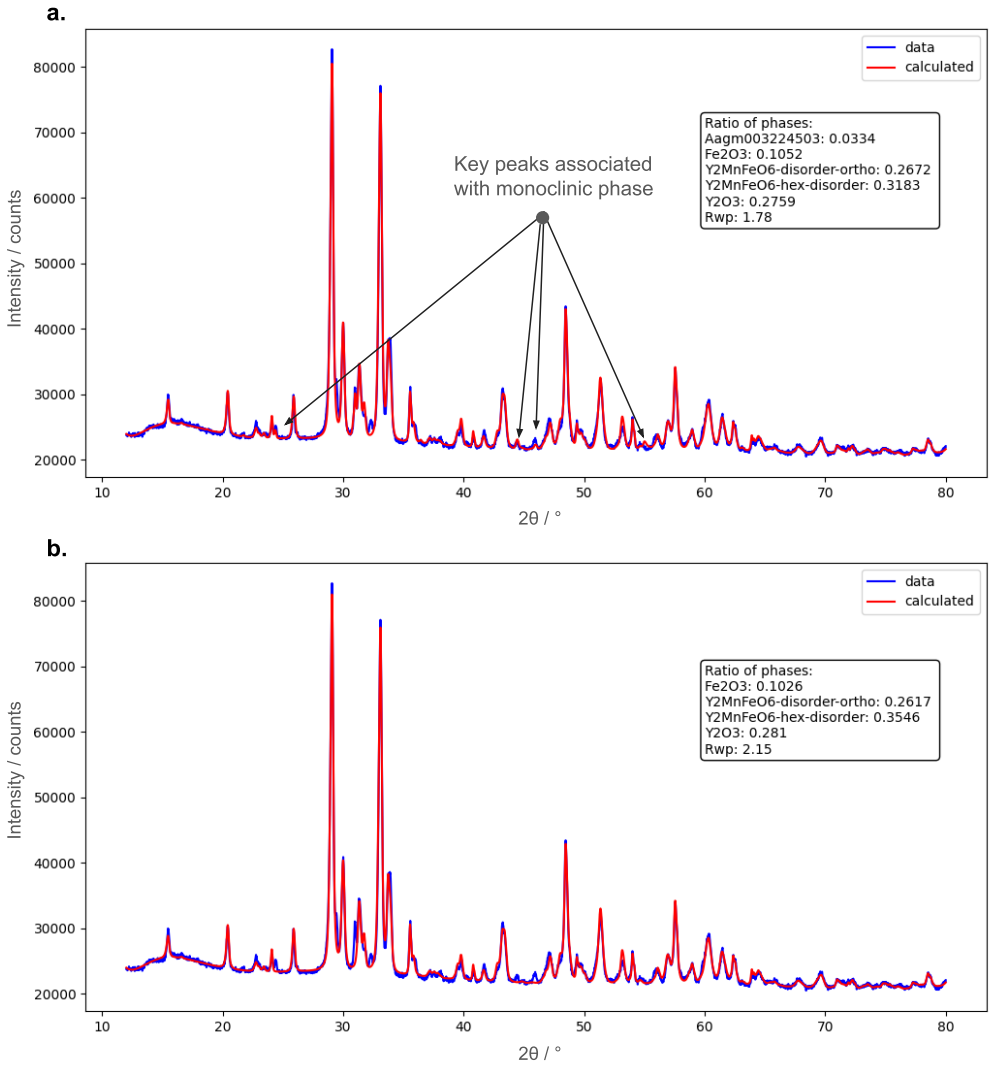

Figure 7: XRD fit for Y_2MnFeO_6. The small inset rectangle contains phases present (Y_2O_3, Fe_2O_3, and several disordered phases of Y_2MnFeO_6. **a.** Fit with the target phase (labeled Aagm003224503) present versus **b.** without the target phase present. A notable improvement in R_{wp} is observed, as well as multiple previously unfitted peaks being fitted, strongly suggesting the presence of the target structure. Known Y-Mn-Fe-O compounds (including two and three metal compounds) did not fit the target peaks.

Eu_2WO_6:

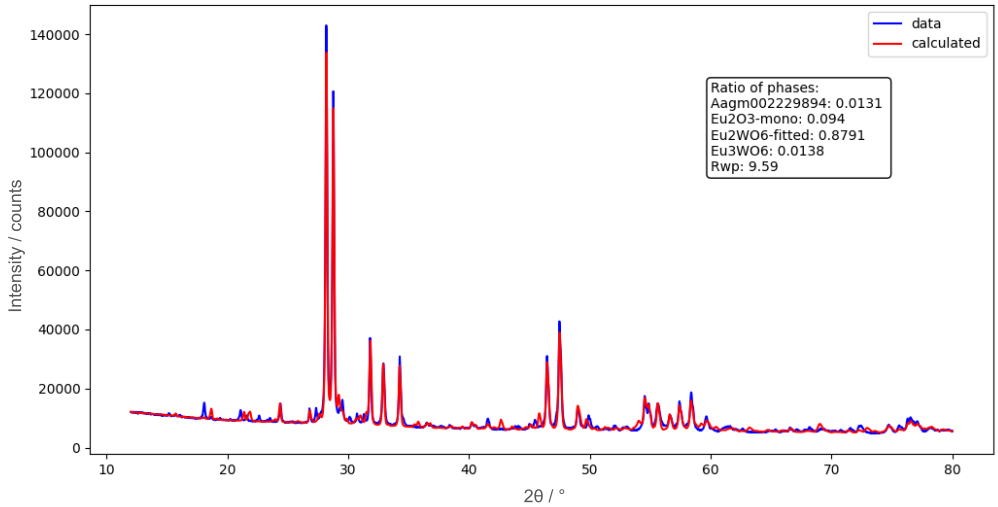

Figure 8: XRD fit for Eu_2WO_6. The target phase is observed at too low a percentage (<1.5% to be plausible. Instead a polymorph (88% purity) with a C2/m with a=14.29 Å, b=3.649 Å, c = 8.876 Å, and $\beta = 100.4°$ unit cell is observed instead. Some intensity, likely a complex tungsten oxide, could not be fitted.

$DyAl_3(BO_3)_4$:

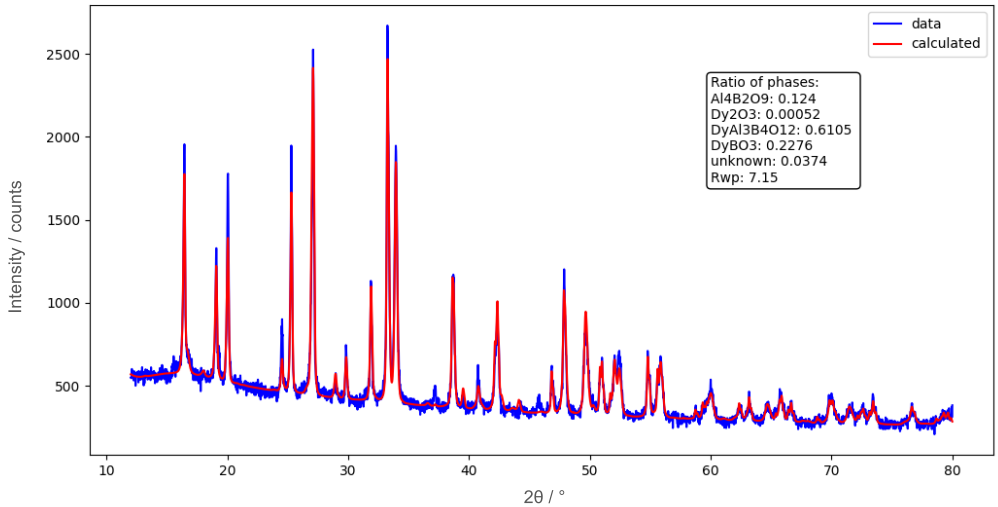

Figure 9: XRD fit for $DyAl_3(BO_3)_4$. The target phase is observed at 62%. An unknown orthorhombic phase that does not correspond to known structures was observed at 4%.

$Nd_3Mn_2Sb_3O_14$:

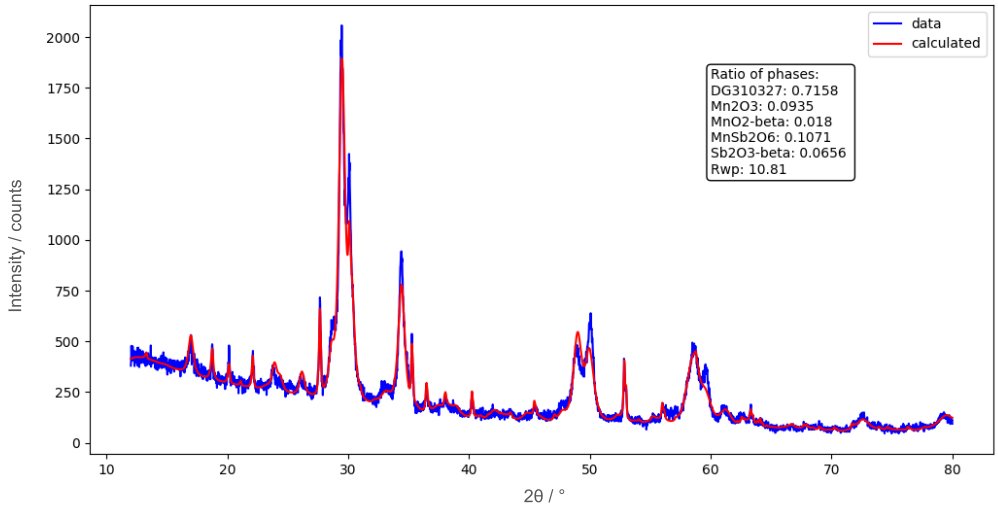

Figure 10: XRD fit for $Nd_3Mn_2Sb_3O_14$. The target phase is listed DG310527 and has 72% purity.

Y_2MnCoO_6:

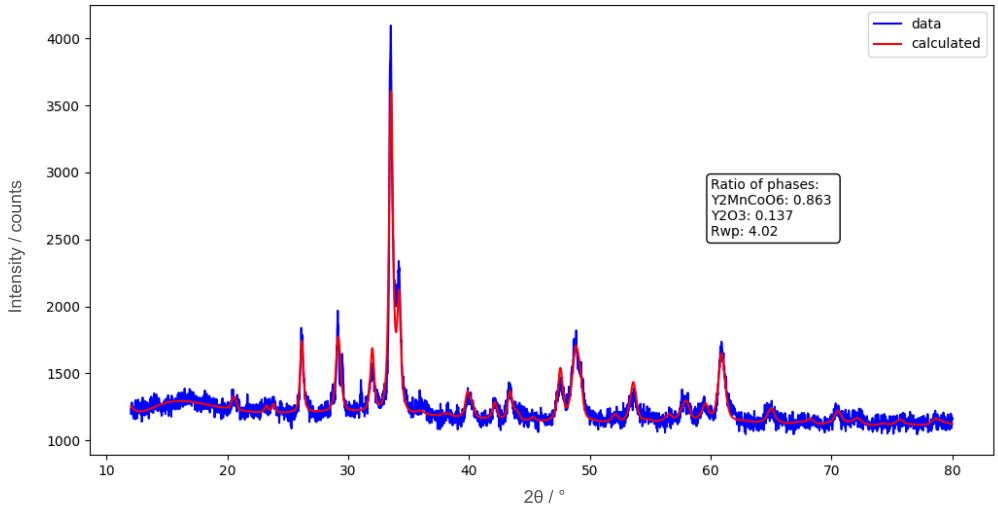

Figure 11: XRD fit for Y_2MnCoO_6.

Sm_3BWO_9:

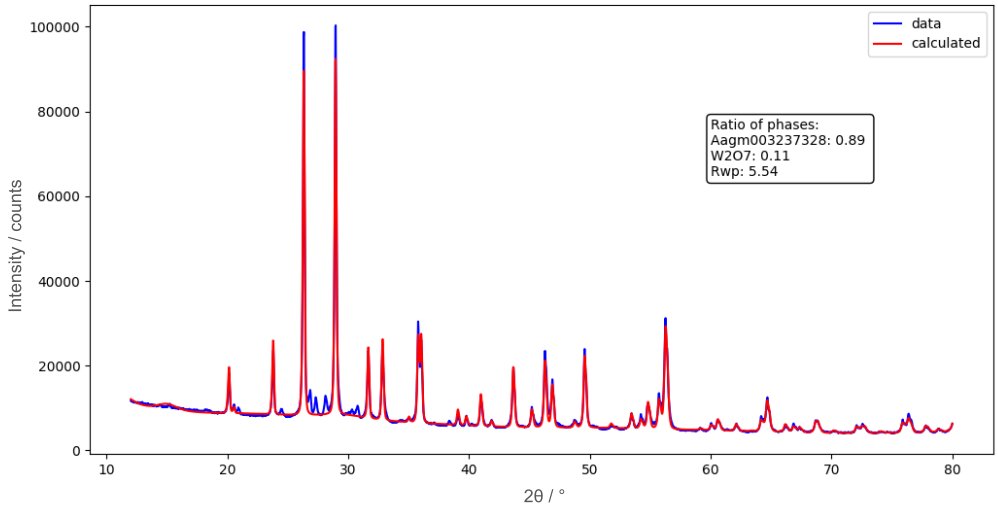

Figure 12: XRD fit for Sm_3BWO_9. The target phase is listed Aagm003237328.

$Sr_4Al_6MoO_16$:

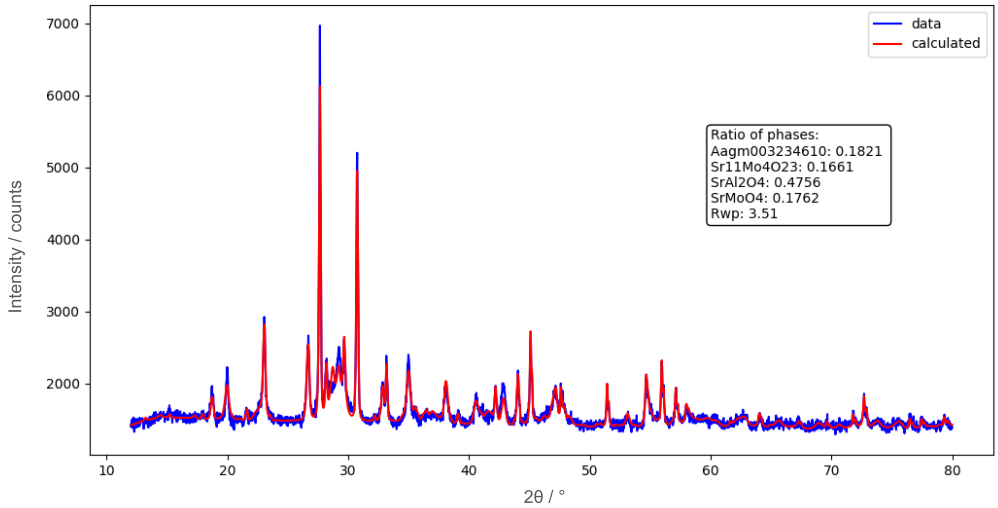

Figure 13: XRD fit for $Sr_4Al_6MoO_{16}$. The target phase is listed Aagm003234610.

$TbFeO_3$:

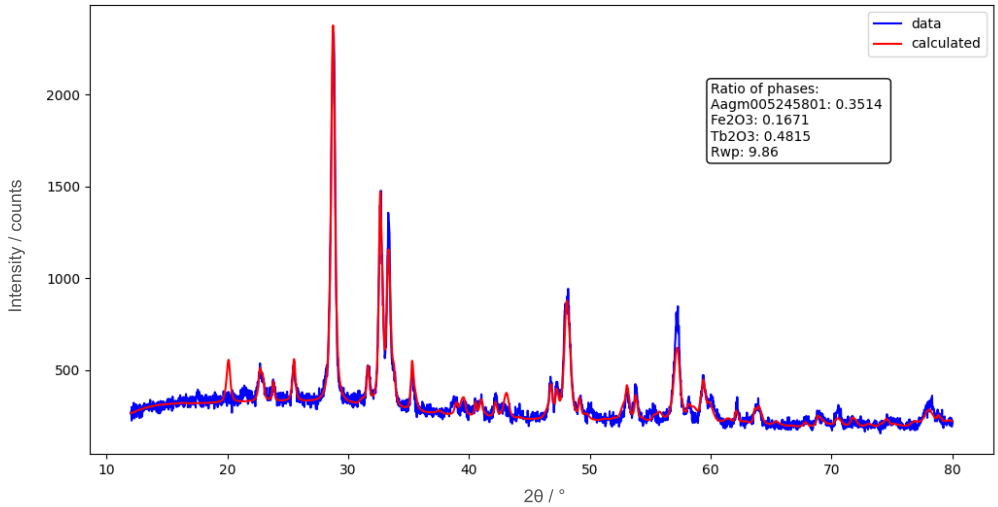

Figure 14: XRD fit for $TbFeO_3$. The target phase is listed Aagm005245801.

$Ba_2Gd(CuO_2)_4$: (provided as an example of a failed synthesis)

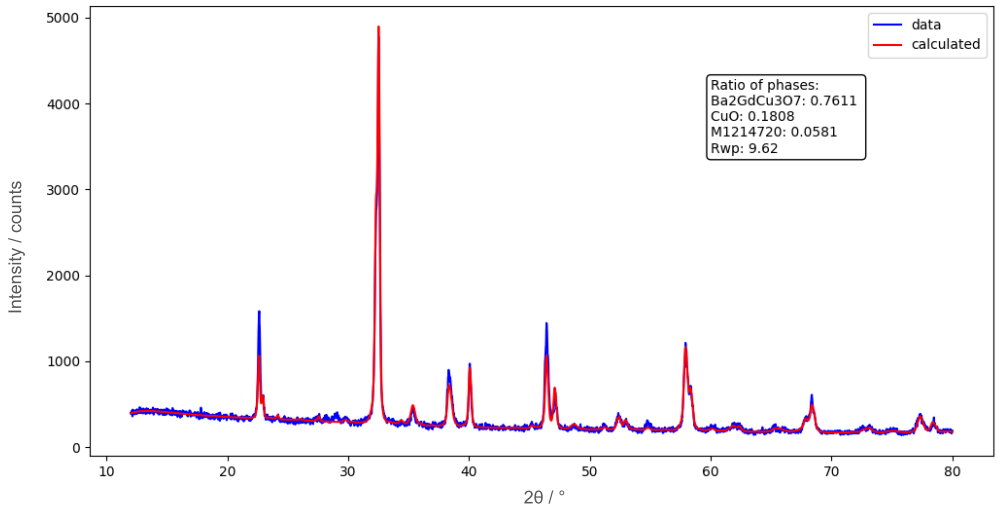

Figure 15: XRD fit for $Ba_2Gd(CuO_2)_4$. $Ba_2GdCu_3O_7$ was made instead.

A.5 CIF File for Nd_3BTeO_9

```
data_Nd3BTeO9
_symmetry_space_group_name_H-M   'P 1'
_cell_length_a   8.76075
_cell_length_b   8.76075
_cell_length_c   5.55533
_cell_angle_alpha   90.000
_cell_angle_beta   90.000
_cell_angle_gamma   120.000
_symmetry_Int_Tables_number   1
_chemical_formula_structural   Nd3BTeO9
_chemical_formula_sum   'Nd6 B2 Te2 O18'
_cell_volume   369.25218
_cell_formula_units_Z   2
loop_
 _symmetry_equiv_pos_site_id
 _symmetry_equiv_pos_as_xyz
  1  'x, y, z'
loop_
 _atom_site_type_symbol
 _atom_site_label
 _atom_site_symmetry_multiplicity
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_occupancy
  Nd  Nd0  1  0.351800  0.072868  0.204851  1
  Nd  Nd1  1  0.927133  0.278934  0.204851  1
  Nd  Nd2  1  0.721067  0.648200  0.204851  1
  Nd  Nd3  1  0.278934  0.351801  0.704850  1
  Nd  Nd4  1  0.072867  0.721067  0.704850  1
  Nd  Nd5  1  0.648200  0.927133  0.704850  1
  B  B6  1  0.000000  0.000000  0.373877  1
  B  B7  1  0.000000  0.000000  0.873876  1
  Te  Te8  1  0.333334  0.666667  0.246297  1
  Te  Te9  1  0.666666  0.333334  0.746298  1
  O  O10  1  0.537430  0.753814  0.045933  1
  O  O11  1  0.246186  0.783615  0.045933  1
  O  O12  1  0.216386  0.462571  0.045935  1
  O  O13  1  0.129442  0.173277  0.365463  1
  O  O14  1  0.043835  0.870558  0.365463  1
  O  O15  1  0.826722  0.956164  0.365463  1
  O  O16  1  0.132779  0.527094  0.461794  1
  O  O17  1  0.472906  0.605685  0.461794  1
  O  O18  1  0.394315  0.867220  0.461794  1
  O  O19  1  0.753814  0.216385  0.545934  1
  O  O20  1  0.462570  0.246185  0.545934  1
  O  O21  1  0.783614  0.537430  0.545934  1
  O  O22  1  0.173277  0.043835  0.865464  1
  O  O23  1  0.956164  0.129442  0.865464  1
  O  O24  1  0.870558  0.826722  0.865464  1
  O  O25  1  0.605684  0.132779  0.961792  1
  O  O26  1  0.867220  0.472905  0.961792  1
  O  O27  1  0.527094  0.394314  0.961794  1
```