

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 TP, FP, TN, 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 :

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

A.2 Additional Synthesizability Prediction Results

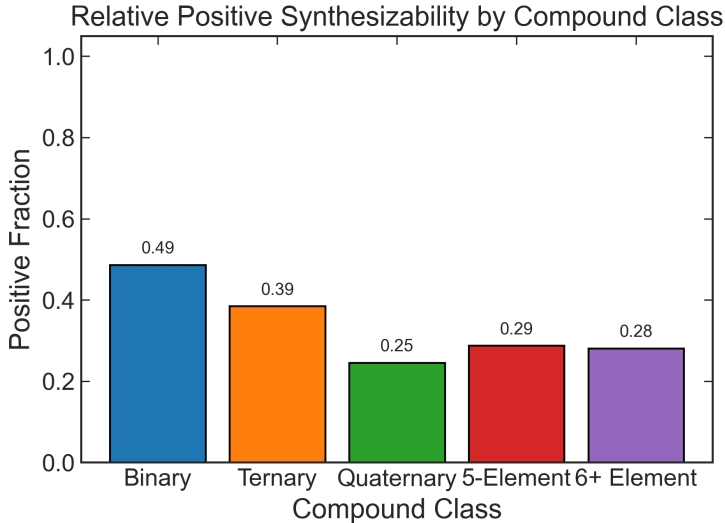


Figure 5: Synthesizability by compound class. Computed from entries in the Materials Project.

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

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

Target phase	Precursors	Temperature	Time
$\text{Y}_2\text{MnFeO}_6^{\text{A}}$	$\text{Y}_2\text{O}_3, \text{Mn}_2\text{O}_3, \text{Fe}_3\text{O}_4$	1030°C	8h
$\text{Nd}_3\text{BTeO}_9^{\text{G}}$	$\text{Nd}_2\text{O}_3, \text{B}_2\text{O}_3, \text{TeO}_2$	965°C	8h
$\text{Eu}_2\text{WO}_6^{\text{A}}$	$\text{Eu}_2\text{O}_3, \text{WO}_3$	1030°C	8h
$\text{DyAl}_3(\text{BO}_3)_4^{\text{A}}$	$\text{Dy}_2\text{O}_3, \text{Al}_2\text{O}_3, \text{B}_2\text{O}_3$	965°C	8h
$\text{Nd}_3\text{Mn}_2\text{Sb}_3\text{O}_{14}^{\text{A}}$	$\text{Nd}_2\text{O}_3, \text{MnO}_2, \text{Sb}_2\text{O}_3$	1030°C	8h
$\text{Y}_2\text{MnCoO}_6^{\text{A}}$	$\text{Y}_2\text{O}_3, \text{Mn}_2\text{O}_3, \text{Co}_3\text{O}_4$	1030°C	8h
$\text{Sm}_3\text{BWO}_9^{\text{A}}$	$\text{Sm}_2\text{O}_3, \text{B}_2\text{O}_3, \text{WO}_3$	965°C	8h
$\text{Sr}_4\text{Al}_6\text{MoO}_{16}^{\text{G}}$	$\text{SrCO}_3, \text{Al}_2\text{O}_3, \text{MoO}_3$	1030°C	8h
$\text{TbFeO}_3^{\text{A}}$	$\text{Tb}_4\text{O}_7, \text{Fe}_2\text{O}_3$	1030°C	8h
$\text{Ba}_2\text{DyFeO}_5^{\text{G}}$	$\text{BaCO}_3, \text{Dy}_2\text{O}_3, \text{Fe}_2\text{O}_3$	965°C	8h
$\text{Ba}_2\text{Gd}(\text{CuO}_2)_4^{\text{M}}$	$\text{BaCO}_3, \text{Gd}_2\text{O}_3, \text{CuO}$	1030°C	8h
$\text{Ba}_2\text{TmFeO}_5^{\text{G}}$	$\text{BaCO}_3, \text{Tm}_2\text{O}_3, \text{Fe}_2\text{O}_3$	965°C	8h
$\text{Eu}_3\text{MoO}_7^{\text{A}}$	$\text{Eu}_2\text{O}_3, \text{MoO}_3$	1030°C	8h
$\text{La}_2\text{Te}_2\text{WO}_{10}^{\text{G}}$	$\text{La}_2\text{O}_3, \text{TeO}_2, \text{WO}_3$	965°C	8h
$\text{Nd}_3\text{BMoO}_9^{\text{G}}$	$\text{Nd}_2\text{O}_3, \text{B}(\text{OH})_3, \text{MoO}_3$	1030°C	8h
$\text{Ho}_2\text{Mn}_2\text{O}_7^{\text{M}}$	$\text{Ho}_2\text{O}_3, \text{MnO}_2$	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 :

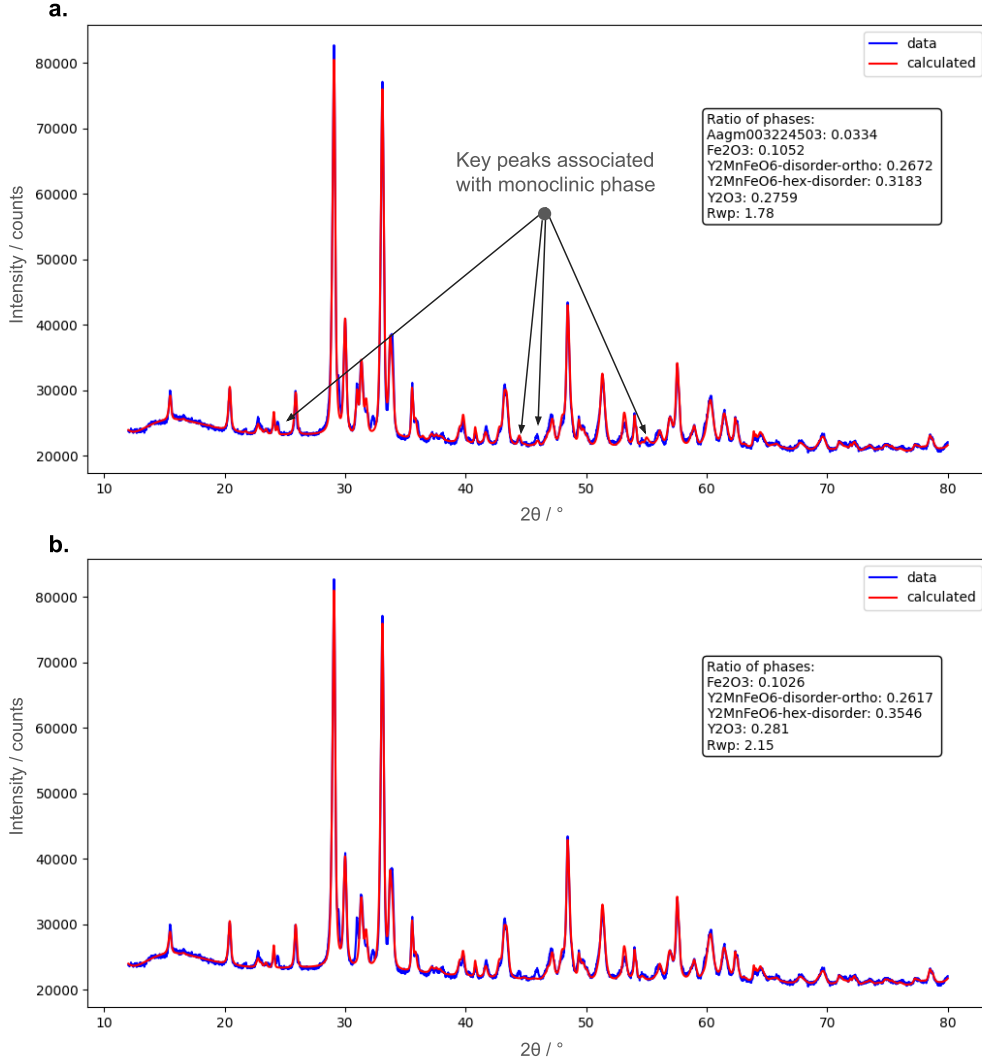


Figure 7: XRD fit for Y_2MnFeO_6 . The small inset rectangle contains phases present (Y_2O_3 , Fe_2O_3 , and several disordered phases of Y_2MnFeO_6). **a.** Fit with the target phase (labeled Aagm003224503) present versus **b.** without the target phase present. A notable improvement in R_{wp} is observed, 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.

Eu₂WO₆:

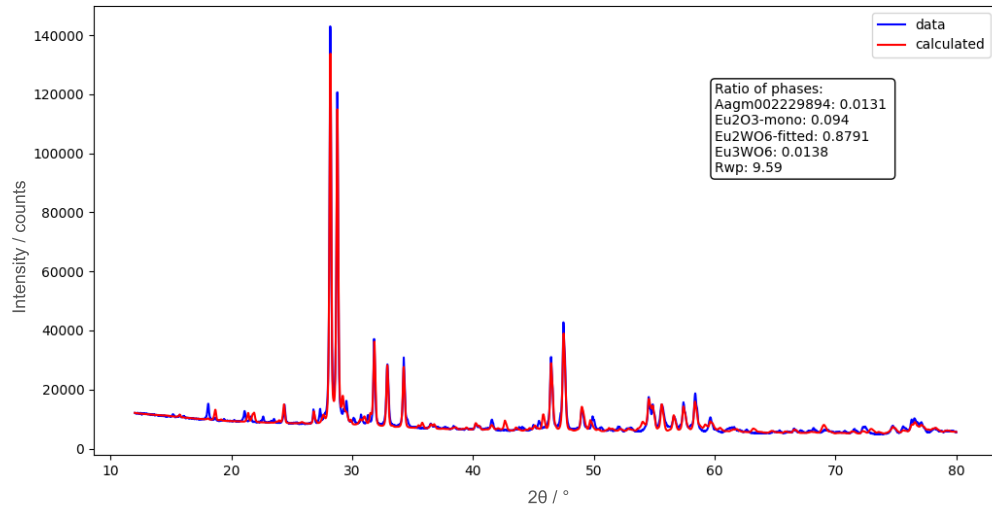


Figure 8: XRD fit for Eu₂WO₆. The target phase is observed at too low a percentage (<1.5% to be plausible). Instead a polymorph (88% purity) with a C2/m with a=14.29 Å, b=3.649 Å, c = 8.876 Å, and $\beta = 100.4^\circ$ unit cell is observed instead. Some intensity, likely a complex tungsten oxide, could not be fitted.

DyAl₃(BO₃)₄:

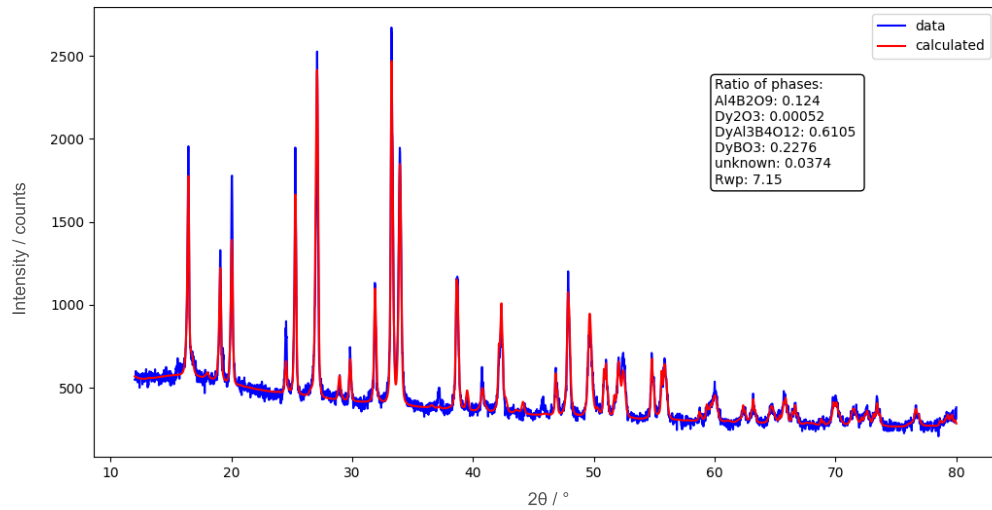


Figure 9: XRD fit for DyAl₃(BO₃)₄. The target phase is observed at 62%. An unknown orthorhombic phase that does not correspond to known structures was observed at 4%.

$\text{Nd}_3\text{Mn}_2\text{Sb}_3\text{O}_{14}$:

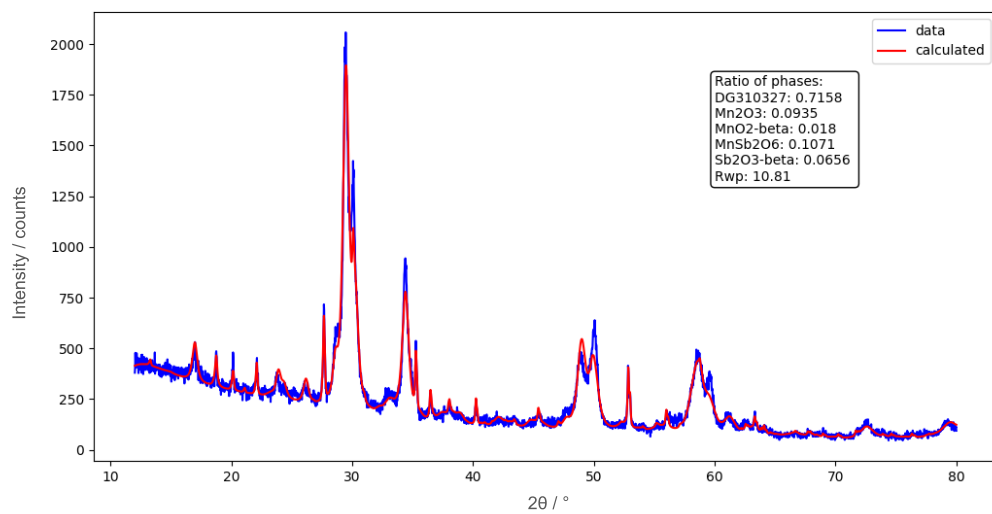


Figure 10: XRD fit for $\text{Nd}_3\text{Mn}_2\text{Sb}_3\text{O}_{14}$. The target phase is listed DG310527 and has 72% purity.

Y_2MnCoO_6 :

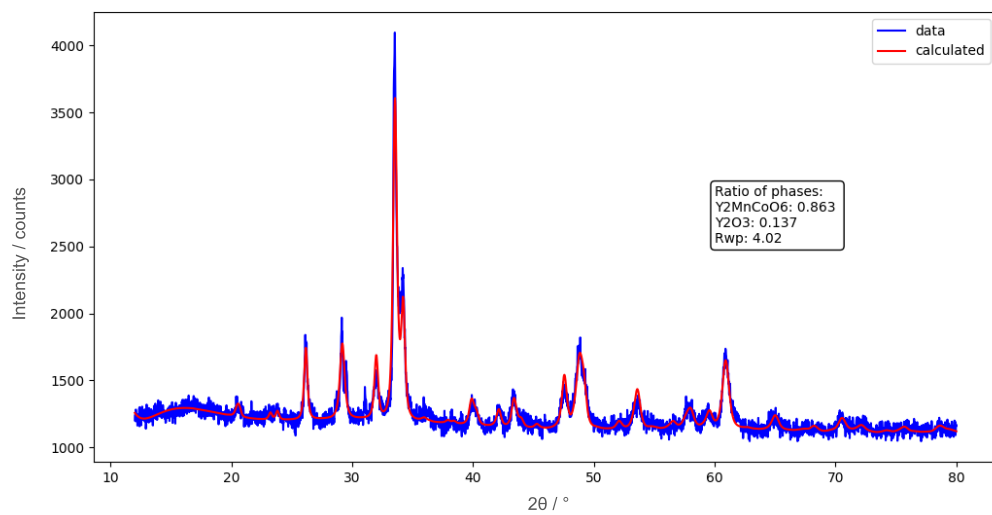


Figure 11: XRD fit for Y_2MnCoO_6 .

Sm₃BWO₉:

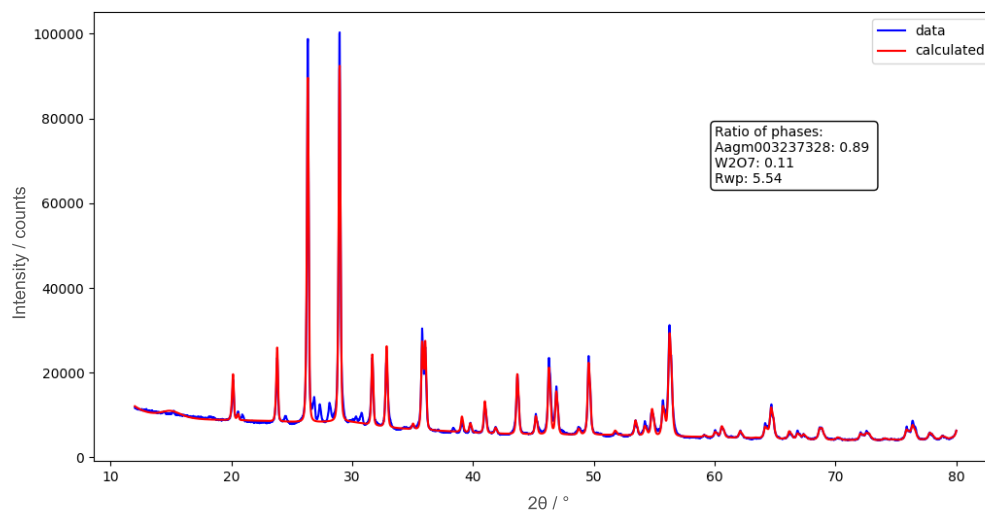


Figure 12: XRD fit for Sm₃BWO₉. The target phase is listed Aagm003237328.

Sr₄Al₆MoO₁₆:

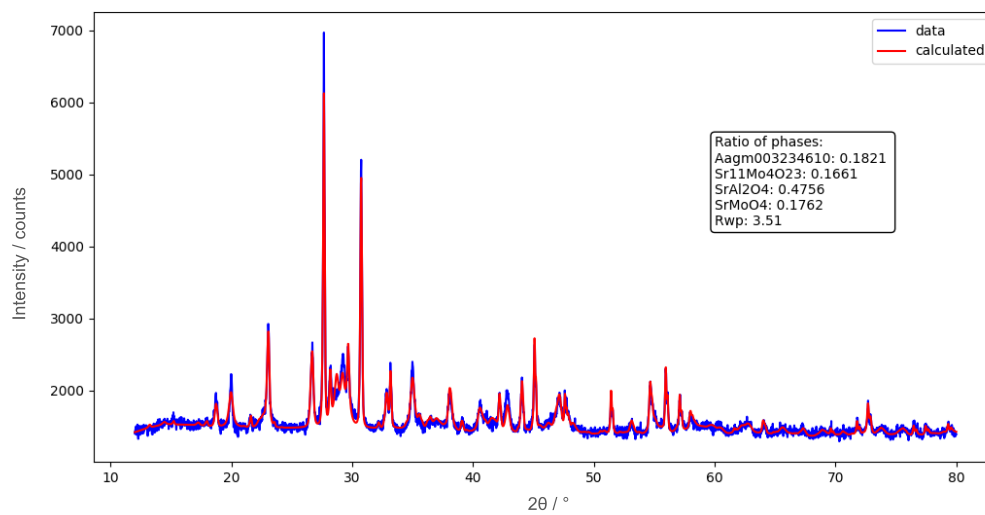


Figure 13: XRD fit for Sr₄Al₆MoO₁₆. The target phase is listed Aagm003234610.

TbFeO₃:

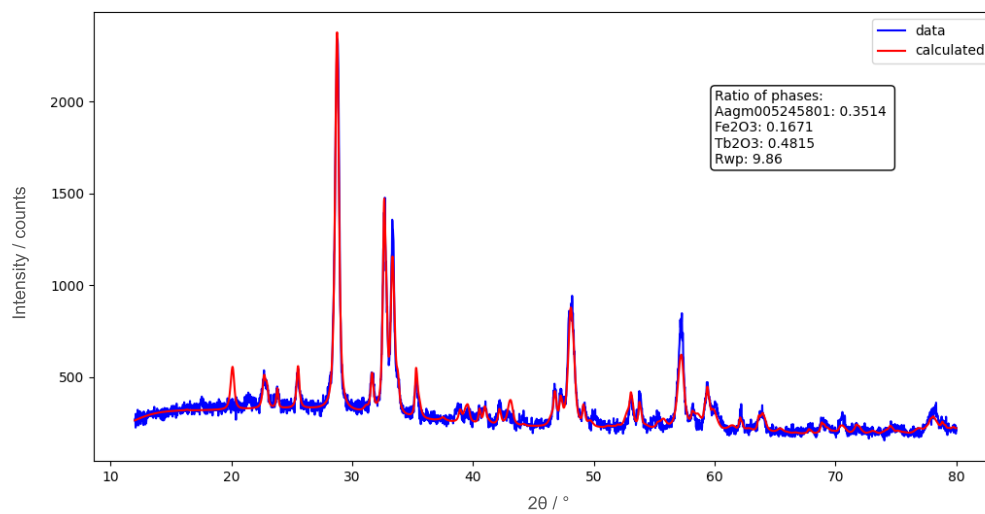


Figure 14: XRD fit for TbFeO₃. The target phase is listed Aagm005245801.

Ba₂Gd(CuO₂)₄: (provided as an example of a failed synthesis)

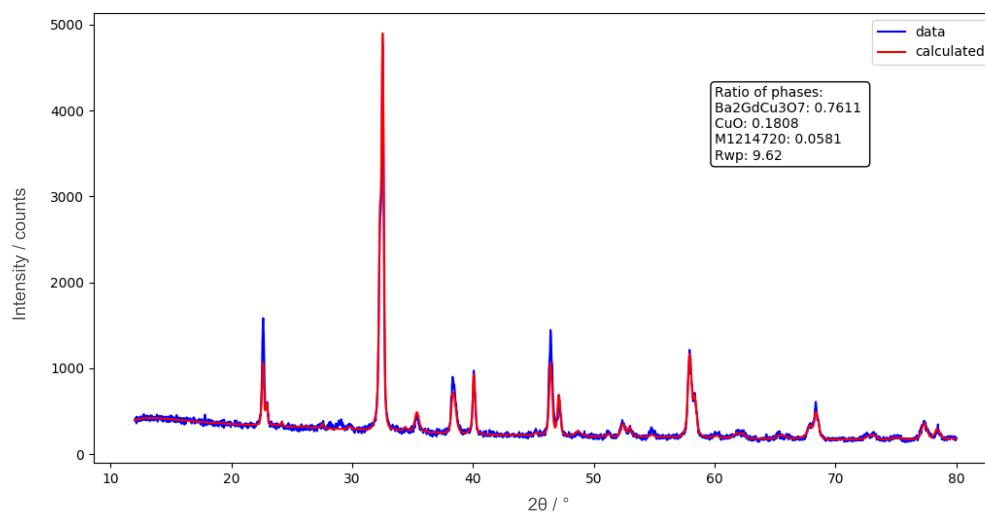


Figure 15: XRD fit for Ba₂Gd(CuO₂)₄. Ba₂GdCu₃O₇ was made instead.

A.5 CIF File for Nd₃BTeO₉

```

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

```