

# Challenges in high-throughput inorganic material prediction and autonomous synthesis

Josh Leeman,<sup>†</sup> Yuhan Liu,<sup>‡</sup> Joseph Stiles,<sup>†,¶</sup> Scott B. Lee,<sup>†</sup> Prajna Bhatt,<sup>‡</sup> Leslie M. Schoop,<sup>\*,†,¶</sup> and Robert G. Palgrave<sup>\*,‡</sup>

<sup>†</sup>*Department of Chemistry, Princeton University, Princeton, NJ 08540, USA*

<sup>‡</sup>*Department of Chemistry, UCL, 20 Gordon St, London, WC1H0AJ, United Kingdom.*

<sup>¶</sup>*Princeton Materials Institute, Princeton University, Princeton, NJ 08540, USA*

E-mail: lschoop@princeton.edu; r.palgrave@ucl.ac.uk

## Abstract

Materials discovery lays the foundation for many technological advancements. Predicting and discovering new materials are not simple tasks. We here outline some basic principles of solid-state chemistry, which might help to advance both, and discuss pitfalls and challenges in materials discovery. Using the recent work of Szymanski et al., which reported the autonomous discovery of 43 novel materials, as an example, we discuss problems that can arise in unsupervised materials discovery, and hope that by addressing these, autonomous materials discovery can be brought closer to reality. We discuss all 43 synthetic products and point out four common shortfalls in the analysis. These errors unfortunately lead to the conclusion that no new materials have been discovered in that work. We conclude that there are two important points of improvement that require future work from the community: (i) automated Rietveld analysis of powder x-ray diffraction data is not yet reliable. Future improvement of such, and the development of a reliable artificial intelligence-based tool for Rietveld

fitting, would be very helpful, not only to autonomous materials discovery, but also the community in general. (ii) We find that disorder in materials is often neglected in predictions. The predicted compounds investigated herein have all their elemental components located on distinct crystallographic positions, but in reality, elements can share crystallographic sites, resulting in higher symmetry space groups and - very often - known alloys or solid solutions. This error might be related to the difficulty of modeling disorder in a computationally economical way, and needs to be addressed both by computational and experimental material scientists. We find that two-thirds of the claimed successful materials in Szymanski et al are likely to be known, compositionally disordered versions of the predicted, ordered compounds. We highlight important issues in materials discovery, computational chemistry, and autonomous interpretation of x-ray diffraction. We discuss concepts of materials discovery from an experimentalist point of view, which we hope will be helpful for the community to further advance this important new aspect of our field.

## Introduction

Inorganic materials serve as the basis of modern technology. This has always been the case, and it is no coincidence that we have named several historical epochs after inorganic materials. Many known crystalline inorganic materials are tabulated in the Inorganic Crystal Structure Database (ICSD),<sup>1</sup> which currently has about 200,000 entries, although not all of those are unique compounds.

Material scientists heavily rely on this database to find materials with relevant properties, which would, for example, improve the current state of the art in Li-ion batteries, make data storage more efficient, increase the efficiency of solar cells, and much more. Expanding the library with new and reliable inorganic materials can help with these endeavors.

Since the development of chemistry as a distinct science, new materials have been discovered in laboratories, either with targeted syntheses, testing new compositions, determining

phase diagrams, or accidentally. More modern methods utilize computation as a guide. Still, the process is tedious and the ICSD expands slowly. The Materials Project,<sup>2</sup> has been one approach to expand the space of inorganic materials. It catalogues the known ICSD compounds with additional computational information and also suggests computationally predicted new materials.

Very recently, Google DeepMind reported the prediction of up to 2.2 million new, stable inorganic crystals, tabulated in their GNoME database.<sup>3</sup> Some of these predictions may warrant experimental verification. Synthesizing so many material candidates by hand would be extremely laborious. To accelerate this, a group at Berkeley established an automated lab, using robotics and artificial intelligence, called A-lab. A-lab uses robots to mix and heat ingredients and measure powder x-ray diffraction (PXRD) of the products. An algorithm then analyzes the PXRD patterns, decides whether the synthesis was successful, and if not, adjusts the synthetic conditions. The group behind A-lab recently reported that within 17 days, A-lab was able synthesize 41 new materials out of 58 predicted targets, an impressive success rate of 71%.<sup>4</sup> Using human intervention the success rate was increased to 78% — 43 successfully synthesized new materials. If this were true, it could drastically accelerate materials discovery, potentially yielding hundreds of new compounds annually. Throughout this comment, we will refer to the work of Szymanski et al. as the “A-lab paper”.

Many aspects of this work are impressive: the fact that robots can take over labor intensive steps, that AI can predict reasonable synthetic routes based on literature precedent, and that a full circle of materials synthesis and characterization without human intervention can be carried out. Unfortunately, we found that the central claim of the A-lab paper, namely that a large number of previously unknown materials were synthesized, does not hold. As we will explain below, we believe that at time of publication, none of the materials produced by A-lab were new: the large majority were misclassified, and a smaller number were correctly identified but already known. In this latter category, three compounds have been reported in between GNoMEs screenshot of the ICSD and time of the A lab publication, meaning

that they would not be in the original training set.

Additionally, we find that the vast majority of the synthetic products were wrongly characterized. These misinterpreted characterizations broadly fall under two categories: either authors failed to recognize that the automated refinement process changed the symmetry of the target compound, or the PXRD pattern agrees better with known, or more often a mixture of known phases. A more detailed explanation of these pitfalls will be laid out in a following section. In general, it seems that one issue lies with the final characterization step (in this case the Rietveld refinement), thus improvement of AI-assisted materials characterization seems to be one the bottlenecks of automated materials discovery. Another might be related to the role of disorder in materials and how this is often not modeled or considered when new materials are predicted. Thus materials prediction could also be improved by considering the role of disorder.

Before we dive into the PXRD data analysis, we first briefly discuss what makes a material ‘new’. As we hope to reach a multi-disciplinary audience, we will then go into some, but not all, standard practices of the field when validating this claim. We detail thematic issues that arise when analyzing A-lab’s data. Addressing these issues would likely make automated lab projects more reliable, and then serve as a more useful tool for solid-state chemists. The bulk of this paper is an analysis of the 43 compounds sorted categorically, both to compartmentalize the reported compounds efficiently and to highlight the types of errors we see as motifs arising in each class.

## **What constitutes a ‘new’ inorganic material?**

Chemists usually distinguish inorganic materials by their structure and composition, and in some cases, properties. The dominance of x-ray crystallography in the study of the solid state commonly leads to delineation between materials based on their diffraction properties, along with analysis of their composition by methods such as atomic emission spectroscopy

or mass spectrometry. Pure molecular materials take their composition from their molecular formula, but may still form different crystal structures, or polymorphs, which have different properties and are often considered to be distinct materials. Non-molecular materials, such as those we are concerned with here, can also show polymorphism, the most famous example being diamond and graphite as polymorphs of carbon, but in addition to this, non-molecular materials do not have a chemical composition restricted by molecular formulae. Their composition is not quantized, but instead can be incrementally changed. One primary example of this is solid solutions, for instance, a solution of KCl and NaCl, which could be written  $\text{Na}_{1-x}\text{K}_x\text{Cl}$ , where  $x$  can take any value between 0 and 1.<sup>5</sup> Doping is a related concept where some percentage of an impurity is incorporated into a material; doped silicon is the basis of modern electronics due to the large effect on electron transport imparted by a small concentration of impurity. This ability to incrementally alter composition challenges concepts of what constitutes a *new material*.

A central theme in our analysis of the work presented in the A-lab paper, which we believe pertains more widely to the field of high-throughput computational material prediction, is the concept of order and disorder within a crystal lattice. The defining characteristic of a crystal lattice is order, but compositional disorder of atoms within a lattice is a widespread phenomenon. In fact, disorder in a crystal lattice is often used to tune the properties of a material, an example of which was recently demonstrated in  $\text{Li}_{1.2}\text{Cr}_{0.4}\text{Mn}_{0.4}\text{O}_2$ .<sup>6</sup> Another example is the aforementioned  $\text{Na}_{1-x}\text{K}_x\text{Cl}$  solid solution, which adopts the rock salt structure, with the Na and K atoms disordered over the cation sites. Physically, there is a statistical distribution of Na and K in the crystal - the probability of finding one particular cation in one particular location is based on the value of  $x$  in  $\text{Na}_{1-x}\text{K}_x\text{Cl}$ . Such a system can be thought of as structurally ordered but compositionally disordered. Experimental crystallographers can accommodate compositional disorder within the framework of the unit cell description of the crystal, by simply stating that a single crystallographic site may be occupied by a mixture of multiple atom types with fractional occupancy. Thus, in crystallography, such a

disordered system is represented with the same unit cell symmetry that would apply if there were only one atom type on the mixed site (i.e. the symmetry of the aristotype), but then specifying a fractional occupancy for some of the atoms. This description of compositionally disordered materials using partial occupancies has several advantages. Firstly, it is commensurate with the experimental diffraction patterns - the PXRD of  $\text{Na}_{1-x}\text{K}_x\text{Cl}$  solid solutions resemble closely those of NaCl and KCl, but with only small shifts in peak positions and intensities, so it would make sense that the unit cell is very similar too. Secondly, fractional occupancies can be used in the structure factor equation to calculate diffraction intensities, this allows quantitative use of a unit cell with fractional occupancies for example in Rietveld refinement, while simple heuristics like Vegard's law relate the composition of a solid solution to the lattice parameter, usually with good accuracy. The usefulness of the idea of fractional occupancy, and its compatibility with many experimental crystallographic methods, is such that it is easy to overlook that in fact it breaks the foundational assumption of crystallography, that of translational symmetry.<sup>7</sup> As we will discuss below, this fact becomes much more significant when computational chemistry calculations are undertaken.

Instead of being compositionally disordered, two types of atoms can instead form ordered arrangements. For example, the zincblende structure is an ordered version of the diamond structure. It can further be expanded to the chalcopyrite structure when the cations are ordered.<sup>8</sup> Chalcopyrite ( $\text{CuFeSe}_2$ ) can be viewed as a doubled zincblende lattice, where the Cu and Fe cations order. Now the ordering of the ions causes the unit cell to enlarge, lowering the symmetry, and changing the space group, with concurrent changes to the diffraction pattern. Another well-known example of such ordered superstructures is the double perovskite structure.<sup>9</sup> In the case of alloys, Heusler alloys are a common example of ordered intermetallic compounds.<sup>10</sup>

Whether a compound has ordered or disordered atoms can often, but not always, be distinguished by XRD. The larger unit cells and/or lower symmetry of ordered compounds may result in additional diffraction peaks, or changes in intensity of peaks. If the ordered ions

have very similar x-ray scattering factors, which are determined by the number of electrons in the ion, then XRD may not be able to detect their ordering, and may not be able to distinguish between a material with compositionally disordered ions and one where the same ions are ordered.

Some of the issues relating to defining a new material are now clear. For a material to be new it must be different to every other material. But different how? Materials with different crystal structures usually have distinct diffraction patterns and therefore be considered by many to be different. Doped materials may have very similar diffraction patterns to the parent material, but their properties may change markedly. Likewise, in the case of solid solutions, if the arrangement of constituents is random on large length scales, would have diffraction patterns intermediate between the end members. The question of whether doped silicon is a different material to undoped silicon, or whether a solid solution with  $x=0.1$  is a different material to one with  $x=0.2$ , may elicit different answers depending on context or field. A claim of a new material should therefore be accompanied by an explanation of how it relates to currently known materials, and what differences in structure and composition, or other factors, distinguish them.

Interestingly, it seems that many of the predicted new materials, both in Materials Project and the larger Google's GNoMe, fall in the category of structurally new materials. We have certainly not looked through all predicted new materials, not even a large fraction. By focusing on those that were picked as synthetic targets in the A-lab paper,<sup>4</sup> we see a clear trend. The predicted new materials can very often be derived from known compounds, in which ions were ordered, rather than fractionally occupied and disordered, within the same arisotype as the known parent. If these ordered structures were synthesized, many indeed would qualify as a new crystallographic compounds. If the key characteristic that distinguishes a new material from a known one is cation order, then the cation order needs to be proven to be real, as otherwise the material would be identical to, or a very similar doped version of, the already known disordered version.

How does A-lab define new materials? They state that they chose targets “from the Materials Project that were marked as ‘theoretical’ (that is, not represented in the ICSD)”. It seems as if the criterion for novelty of a material is its absence from the ICSD. This criterion is open to criticism (for example many known compounds are not in ICSD, especially disordered ones) but nonetheless, we will mostly use this criterion to assess the novelty of the A-lab synthesis products. As it seems as if the A-lab paper set the standard of a new material to be new crystallographic compound - not a new disordered version of a known one - we will test the claims to this standard too.

## How to prove that a synthesized material is new

With ideas of what defines a new material in mind, we can now consider the evidence necessary to determine if one has been produced. The creation and testing of hypotheses is a fundamental feature of science. The best strategy for testing a hypothesis can depend on the context; a *positive testing strategy* is one that looks only for evidence that confirms the hypothesis. It can be appropriate in some special situations where only one working hypothesis exists, but in general it is undesirable, and inappropriate adoption of this strategy is a well-known cognitive bias.<sup>11</sup> Given the large number of materials now known, any hypothesis about discovery of a new material cannot be tested solely by confirmation, but must be also tested against falsification - *i.e.* tests should be carried out to determine if the sample under investigation is instead a known material. Any known material that might realistically form under the synthesis conditions should be considered as a candidate for such testing. For example, if a synthesis is carried out using three elements, X, Y and Z, with the intention to form the ternary compound XYZ, it is prudent to assess whether the diffraction pattern (or any other analysis) can instead be explained by known compounds that can be formed by the reactants, *e.g.* the binary compounds XY, XZ and YZ, or other ternary compounds, like XY<sub>2</sub>Z<sub>4</sub>. Likewise, the unreacted starting materials should also be eliminated from enquiry. If

known materials can adequately explain the experimental evidence, then there is no need to conclude that new materials have been formed. This is a statement of Occam's Razor, which was also expressed in similar terms by Russell: "*Whenever possible, substitute constructions out of known entities for inferences to unknown entities.*"<sup>12</sup>

A positive testing strategy, one that only looks for confirmation of the hypothesis of the presence of a new material, and does not look for alternative explanations involving known materials, is inadequate in a field as well established and densely populated as that of materials chemistry. Instead, any report of a new material must be accompanied by an explanation as to why the experimental evidence is *better* explained by a new material, compared to one or more known materials. In some instances, powder x-ray diffraction might not be capable of providing evidence that can differentiate two materials. Even under perfect experimental conditions, there is information loss by the nature of the PXRD experiment, and this means that there is no one-to-one correspondence between PXRD pattern and structure, so that many theoretical structures may give identical diffraction patterns. Therefore, even if an excellent match between model and experimental XRD can be achieved, this still does not guarantee the modelled compound is the correct one. Schlesinger et al.<sup>13</sup> point out that the "*...mere existence of a plausible crystal structure, a good Rietveld fit with a smooth difference plot, acceptable R-values and a successful checkCIF test does not justify the attribute 'correct structure'.*"

If two candidate materials have very similar diffraction patterns, it may be that PXRD cannot distinguish them, and other techniques must be employed to prove which has been made. When fitting a PXRD pattern, just as when fitting any other data, the most reasonable fit is achieved when the number of fitting parameters is kept as low as possible. In a crystallographic setting, lowering the symmetry of the space group will increase the number of variables in the fit. For this reason, it should be ensured that the improvement of a fit is meaningful when symmetry is lowered. Should the quality of a fit in a high symmetry model be comparable to that of a lower symmetry model, one should pick the former one, again in

accordance with Occam's razor.

Many learned societies, such as the American Chemical Society, the German Chemical Society, and the Royal Society of Chemistry (UK), require not only structural but also compositional information on newly reported materials. Several techniques are available to the solid state chemist: Energy dispersive x-ray spectroscopy (EDX), inductively coupled plasma optical emission spectroscopy (ICP-OES), x-ray wavelength dispersive spectroscopy (WDS), x-ray fluorescence spectroscopy (XRF), x-ray photoelectron spectroscopy (XPS), and electron energy loss spectroscopy (EELS) are well known examples. Use of any of these techniques will, however, usually yield the average elemental composition over a large volume of the sample. This is accurate as a measure of material composition if the sample in question is one pure material, but if the sample is a mixture of materials, then the composition analysis will return an average, which may be unrepresentative of the specific material under investigation. In solid state chemistry, this can be a significant problem, as separation of mixtures is much more challenging than in, for example, solution phase chemistry. If a solid state reaction produces a mixture, and none of the components can be easily dissolved, sublimed, or otherwise removed (a very common scenario with oxide chemistry), it might be challenging to accurately measure the elemental composition of the target material. Thus, synthesis of highly phase-pure samples is normally an important part of new materials discovery, as this is the best route to accurate compositional information. Naturally, phase pure samples have many other advantages when it comes to measuring functional materials properties. The standard set by the A-lab paper of >50% purity being 'success' is therefore anomalous in the usual practice of solid state chemistry.

## **Analysis of the A lab dataset**

Below we go through the materials that have been claimed to be successfully synthesized in the A-lab paper.<sup>4</sup> We summarise our finding here, before going through many examples

in detail. The classification of samples by the A lab themselves is as follows. There are 58 compounds mentioned in total. Of those, 15 are classified as failure, seven as partial success (meaning less than 50 wt% in the final product) and 36 were ‘successes’ (including two that were successful offline, meaning with human intervention in the synthesis).

Within the 36 samples classified as successes, we found that the analysis presented for 35 of them suffered from one or more of the error types described below.

1. **Very poor and obviously incorrect fits.** This means models that are such poor fits to the data, often missing intense diffraction peaks, that they cannot be relied upon either for proof of the structure of the compounds, nor their purity. The poor fitting leads to the inability to identify impurity phases. Since the authors aim to have >50 wt% of their product, it is important to identify what other materials are present in order to assess if the 50% threshold has been met. Additionally, the presence of unreacted starting materials is symptomatic of an incomplete reaction and incorrect reaction conditions. This error type is present in 18/36 compounds.
2. **Using different structures for refinement than were claimed in the paper.** In several cases the CIF supplied in the SI is not the same structure (or composition) as that claimed in the main paper. In several examples even the space group between the two differs. An example is  $\text{Mg}_3\text{NiO}_4$  which we discuss below. This error is present in 8/36 compounds.
3. **No evidence for cation ordering.** The most common error is prediction of compounds which are ordered versions of known disordered compounds. For example, as we will show in detail below, the existence of  $\text{MgTi}_2\text{NiO}_6$  is claimed, which is the same as the known ilmenite structure of the same composition, but the predicted structure has ordered Mg and Ni cations, whereas the known structure has those cations disordered. However, no consideration is given by the authors to the possibility that they may have in fact made the known disordered compound instead of their intended com-

pound. We show below that this is in fact the most likely situation. This error type is present in 24/36 compounds.

4. **Reporting existing compounds as new.** In several cases the claimed new compounds are in fact already reported in the ICSD. This error type is present in 3/36 compounds.

Below, we discuss the 43 materials (which includes the partial successes), going into detail in many cases, to highlight the consistencies of the errors described above. We group the materials by structure type for this discussion. For the analysis, the original published experimental XRD patterns were obtained by digitalizing the data provided in the A-lab paper supplementary information using GetData Graph Digitalizer. Because software had trouble identifying the green dots which represent observed XRD data, the experimental data was obtained by combining the calculated fit with the fit's residual. To align the x-axis values for combining, the acquired XY data were then interpolated the data in Origin. This process is certainly not ideal and yields data of lower quality than the original. Nevertheless, we found it was possible to carry out Rietveld refinement on these datasets. This was carried out in GSAS II, the software used by A-lab. We do not have the experimental parameters for the original data collection, and so peak profiles were determined empirically. We do not claim our fits are definitive or cannot be improved upon, but we highlight in each case the features that make us believe the fits we propose are superior to those provided in the original paper.

Table 1: The distribution of errors in the 36 claimed ‘successful’ syntheses. The X symbol denotes the error is present. Error 1 is a very poor fit, such that the fitted model is meaningless. Error 2 is where a different CIF was used for refinement compared with that in the paper, and on Materials Project. Error 3 is where the predicted structure has ordered cations but there is no evidence for order, and a known, disordered version of the compound exists. Error 4 is where the compound is correctly identified, but is already reported.

Claimed Phases	1	2	3	4	Claimed Phases	1	2	3	4
Ba <sub>2</sub> ZrSnO <sub>6</sub>	X	X	X		Mg <sub>3</sub> MnNi <sub>3</sub> O <sub>8</sub>	X		X	
Ba <sub>6</sub> Na <sub>2</sub> Ta <sub>2</sub> V <sub>2</sub> O <sub>17</sub>	X		X		Mg <sub>3</sub> NiO <sub>4</sub>		X	X	
Ba <sub>6</sub> Na <sub>2</sub> V <sub>2</sub> Sb <sub>2</sub> O <sub>17</sub>	X				MgCuP <sub>2</sub> O <sub>7</sub>		X	X	
CaCo(PO <sub>3</sub> ) <sub>4</sub>			X		MgNi(PO <sub>3</sub> ) <sub>4</sub>	X		X	
CaFe <sub>2</sub> P <sub>2</sub> O <sub>9</sub>					MgTi <sub>2</sub> NiO <sub>6</sub>			X	
CaMn(PO <sub>3</sub> ) <sub>4</sub>			X		MgTi <sub>4</sub> (PO <sub>4</sub> ) <sub>6</sub>				X
CaNi(PO <sub>3</sub> ) <sub>4</sub>			X		MgV <sub>4</sub> Cu <sub>3</sub> O <sub>14</sub>	X	X	X	
FeSb <sub>3</sub> Pb <sub>4</sub> O <sub>13</sub>			X		Mn <sub>2</sub> VPO <sub>7</sub>	X		X	
Hf <sub>2</sub> Sb <sub>2</sub> Pb <sub>4</sub> O <sub>13</sub>			X		Mn <sub>4</sub> Zn <sub>3</sub> (NiO <sub>6</sub> ) <sub>2</sub>			X	
InSb <sub>3</sub> Pb <sub>4</sub> O <sub>13</sub>			X		MnAgO <sub>2</sub>	X			X
K <sub>2</sub> TiCr(PO <sub>4</sub> ) <sub>3</sub>			X		Na <sub>3</sub> Ca <sub>18</sub> Fe(PO <sub>4</sub> ) <sub>14</sub>	X			
K <sub>4</sub> MgFe <sub>3</sub> (PO <sub>4</sub> ) <sub>5</sub>	X				Na <sub>7</sub> Mg <sub>7</sub> Fe <sub>5</sub> (PO <sub>4</sub> ) <sub>12</sub>	X			
K <sub>4</sub> TiSn <sub>3</sub> (PO <sub>5</sub> ) <sub>4</sub>	X				NaCaMgFe(SiO <sub>3</sub> ) <sub>4</sub>		X	X	
KBaPrWO <sub>6</sub>	X				NaMnFe(PO <sub>4</sub> ) <sub>2</sub>	X			
KMn <sub>3</sub> O <sub>6</sub>	X	X	X		Sn <sub>2</sub> Sb <sub>2</sub> Pb <sub>4</sub> O <sub>13</sub>			X	
KNaP <sub>6</sub> (PbO <sub>3</sub> ) <sub>8</sub>	X	X	X		Y <sub>3</sub> In <sub>2</sub> Ga <sub>3</sub> O <sub>12</sub>	X			X
KNaTi <sub>2</sub> (PO <sub>5</sub> ) <sub>2</sub>			X		Zn <sub>2</sub> Cr <sub>3</sub> FeO <sub>8</sub>			X	
KPr <sub>9</sub> (Si <sub>3</sub> O <sub>13</sub> ) <sub>2</sub>	X	X			Zr <sub>2</sub> Sb <sub>2</sub> Pb <sub>4</sub> O <sub>13</sub>			X	

## Rock-salt structured materials

As mentioned in the introduction, the rock-salt or NaCl structure, can host solid solutions when different cations or anions are mixed on their respective sites. In this case the space group does not change and the structure type remains rock-salt. Should those cations order however, both the space group and the structure type would change.

$\text{Mg}_3\text{NiO}_4$  (mp-1099253) is predicted to exist in the primitive cubic space group  $Pm\bar{3}m$ . The predicted structure can be viewed as a rock salt structure where the cation order breaks the F centered lattice (see Fig. 1 (a)). A very similar composition,  $\text{MgNiO}_2$  has been reported to exist in a rock salt structure (space group  $Fm\bar{3}m$ , ICSD entry # 290603), where no cation order was observed (see Fig. 1 (b)) The sample synthesized by A-lab with the composition  $\text{Mg}_3\text{NiO}_4$  was claimed to be a successful synthesis in space group  $Pm\bar{3}m$ . However, the provided structure file, which can be found in the Supplemental Information of the A-lab paper and is shown in Fig. 1 (c), has disordered cations, hence the CIF used for fitting by A-Lab has the space group  $Fm\bar{3}m$ . The powder x-ray can be relatively well indexed with space group  $Fm\bar{3}m$  (as shown in the supplemental files in the paper and also in Fig. 1 (c)) but lacks additional peaks (e.g. the 100 peaks around  $21.1^\circ$ ) that would appear in the original proposed space group  $Pm\bar{3}m$  (see Fig. 1). In fact, the powder pattern and relevant systematic absence conditions of  $h + k$ ,  $h + l$ , and  $k + l = 2n$  agrees very well with the known compound  $\text{MgNiO}_2$  as shown in Fig. 1, however the peaks appear at slightly different diffraction angles, which may suggest a doped material or a solid solution. MgO-NiO solid solutions are well studied as catalytic materials, and solid solutions can be formed across the composition range.<sup>14</sup>  $\text{Mg}_2\text{Ni}_2\text{O}_4$  is reported with lattice parameter  $4.1889(1) \text{ \AA}$ ,<sup>15</sup> whereas the CIF of  $\text{Mg}_3\text{NiO}_4$  provided in the A-lab paper indicates a slightly larger lattice parameter of  $4.20311 \text{ \AA}$ . By interpolating between the lattice parameters of rock salt MgO ( $4.214 \text{ \AA}$ ),<sup>16</sup> and (metastable) rock salt NiO ( $4.1718 \text{ \AA}$ ),<sup>17</sup> Vegard's law places the composition of a solid solution with the lattice parameter  $4.20311 \text{ \AA}$  at exactly  $\text{Mg}_3\text{NiO}_4$ , in line with the expected composition from the synthesis recipe. Note that in the image of the refinement provided in

the A-lab paper, the indexed peaks (tick marks) do not line up with the diffraction peaks, thus we believe there has been an error in producing the image in this case. We conclude that the synthesised compound is actually a member of the MgO-NiO solid solution series, with disordered cations, that has been studied for many years, and not the cation ordered material predicted by Materials Project. Nonetheless, this does not mean that the proposed ordered material cannot be synthesized. However, it will require a different synthetic route to potentially stabilize  $\text{Mg}_3\text{NiO}_4$  with ordered cations. Here, the clear distinction between the PXRD of the ordered and disordered material allows for easy identification of the former. The analysis of  $\text{Mg}_3\text{NiO}_4$  suffers from errors two (different structure used in refinement than was predicted) and three (no evidence for cation order in the predicted structure).

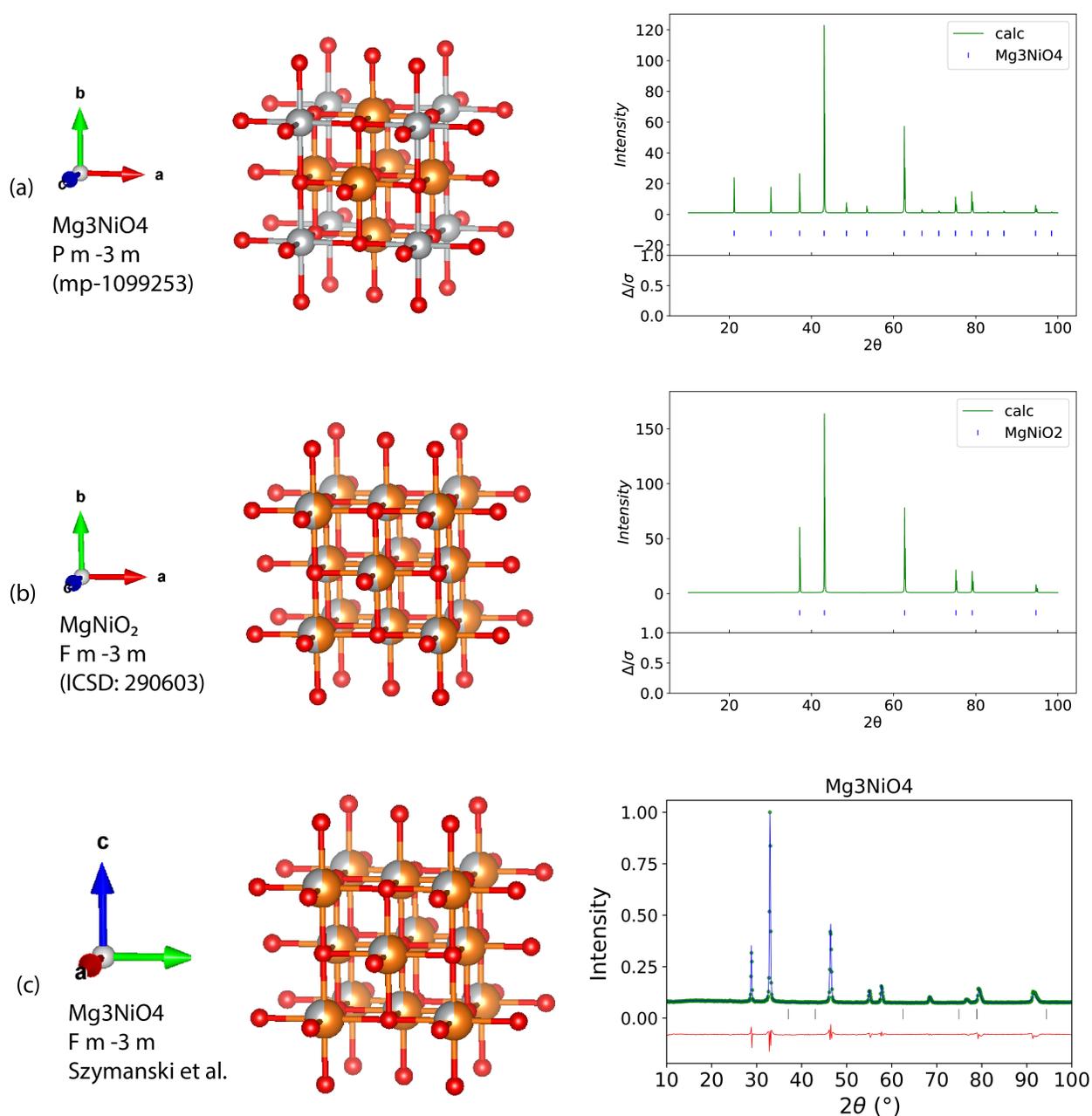


Figure 1: (a) Structure of Mg<sub>3</sub>NiO<sub>4</sub> as predicted by Materials Project (left) and simulated PXR pattern of the same structure (right). (b) Structure of Mg<sub>2</sub>Ni<sub>2</sub>O<sub>4</sub> as listed in the ICSD (left) and simulated PXR pattern of the same structure (right). (c) Structure of Mg<sub>3</sub>NiO<sub>4</sub> as provided by Szymanski et. al (left) and measured powder pattern given in the same paper (right). Mg is shown in orange, Ni in grey and O in red.

Mg<sub>3</sub>MnNi<sub>3</sub>O<sub>8</sub> is predicted by A-lab to exist in the  $R\bar{3}m$  space group. A compound with

the same composition exists in the ICSD, reported by Taguchi et al. in 1995.<sup>18</sup> The reported compound is cubic, a variant of the rock salt structure, sometimes called Murdochite, with octahedrally coordinated metals; the Mn ions form a fcc arrangement, while the Mg and Ni ions, are disordered on a different site. There is also an additional cation vacancy compared with the parent rock salt structure. The A-lab structure, is exactly the same, except the Mg and Ni ions are now ordered, and the particular ordering reduces the symmetry to  $R\bar{3}m$ . The largest effect on the calculated diffraction pattern of this ordering is an increase in the intensity of the peak around  $18.4^\circ$ . In the Murdochite phase ( $Fm\bar{3}m$ ), the (111) peak at  $18.4^\circ$  has an intensity of 33% of the most intense reflection, whereas ordering of the Mg and Ni ions as increases the intensity of this peak to 59% of the most intense reflection. This difference should be easily detectable by the PXRD methods used. Turning to the reported PXRD pattern and refinement by A-lab, it is clear that many of the intensities from the model are very poor matches to the experimental data. Most obviously, the model greatly overestimates the intensity of the reflection at  $18.4^\circ$ . This suggests that the predicted ordering is not present. The generally poor agreement in intensities may point to multiple phases present in this sample. Simple rock salt oxides, such as MgO and NiO, have intense peaks that coincide with some of the Murdochite peaks (unsurprising as they are based on the same structural motif), and so the incorrect intensities may be due to the presence of rock salt phases. There is also an almost completely unmodelled peak in the experimental pattern at just over  $30^\circ$ . This peak is not present in the disordered  $Mg_3MnNi_3O_8$ , nor is it a rock salt (MgO or NiO) peak. It is, however, present at reasonable intensity in the pattern of  $NiMn_2O_4$  spinel, and to us this (or a similar spinel) seems the best candidate to explain that peak. The sample therefore may consist of multiple phases: rock salt, spinel, Murdochite, with the Mg, Mn, Ni possibly distributed across all these phases. The evidence from the peak intensities is clearly against the proposed Ni-Mg ordering. The analysis of  $Mg_3MnNi_3O_8$  suffers from error one (very poor fit) and error three (no evidence of cation order).

## Layered Materials

Layered materials are of significant interest in materials science, as they provide the foundation of many applications, including most battery electrode materials.<sup>19</sup> Among the 43 materials that A-lab synthesized, there is one layered compound,  $\text{KMn}_3\text{O}_6$ . This compound was predicted by Materials Project (mp-1016190) to crystallize in space group  $C2/c$ . The structure can be viewed as related to  $\alpha\text{-NaFeO}_2$ , which consists of layers of edge-sharing  $\text{FeO}_6$  octahedra with Na cations between the layers and crystallized in space group  $R\bar{3}m$ . Variations of this structure exist in several space groups, where the layer stacking causes symmetry change.<sup>20</sup> Cation order in the transition metal layer can lower the symmetry to the monoclinic space group  $C2/m$ .<sup>21</sup> The proposed structure of  $\text{KMn}_3\text{O}_6$  does not have cation order on the transition metal site, but proposes ordered vacancies of K (Fig. 2(a)). In contrast, the structure that is reported in the SI of the A-lab paper<sup>4</sup> has disordered K and the actual space group of the provided structure is  $C2/m$ , not  $C2/c$  (Fig. 2(b)). This is still low symmetry for a material, which might be better described as  $\text{K}_{0.33}\text{MnO}_2$ . Intuitively, one would expect  $\text{K}_{0.33}\text{MnO}_2$  would adopt one of the  $\alpha\text{-NaFeO}_2$ -structure variants, which usually have hexagonal or rhombohedral symmetry. The low symmetry in the predicted material likely arises from the slight buckling of the layers, which can be seen in Fig. 2 (a) and (b). Both the K order as well the buckling of the layers would likely define this material to be new, but at this point it is unclear if either of those are present in the synthetic product described in the A-lab paper.<sup>4</sup> For example,  $\text{K}_{0.3}\text{MnO}_2$  has been reported in the hexagonal space group  $P6_3/mmc$ , in a structure that belongs to one of the stacking variants of the  $\alpha\text{-NaFeO}_2$ -structure.<sup>22</sup> This structure is shown in Fig. 2 (c); it has neither ordered K nor buckled layers.  $\text{K}_{0.3}\text{MnO}_2$  is known to result from the thermal decomposition of  $\text{KMnO}_4$  above  $800^\circ\text{C}$ .<sup>22</sup> As A-lab's reaction conditions included a  $1000^\circ\text{C}$  heating step,  $\text{K}_{0.3}\text{MnO}_2$  is a likely product. The PXRD fit for  $\text{KMn}_3\text{O}_6$  provided in the A-lab paper<sup>4</sup> is of very poor quality and misses some major reflections as shown in Fig. 2 (d). Comparing it to the simulated PXRD pattern of  $\text{K}_{0.3}\text{MnO}_2$  (Fig. 2 (e)), reveals that the main measured reflections are well

reproduced by those simulated to appear  $K_{0.3}MnO_2$ . Still, there is an intensity mismatch, which could be caused by preferred orientating, impurities, or additional phases. There are many layered K-Mn-O materials in the literature that could also explain the PXRD data. Thus, proof for the proposed structure is lacking and the more likely explanation for the synthetic product is one, or a combination of, known layered K-Mn-O phases. The analysis of  $KMn_3O_6$  suffers from error one (poor fit), two (CIF file not the same as originally predicted), and three (so evidence of cation order).

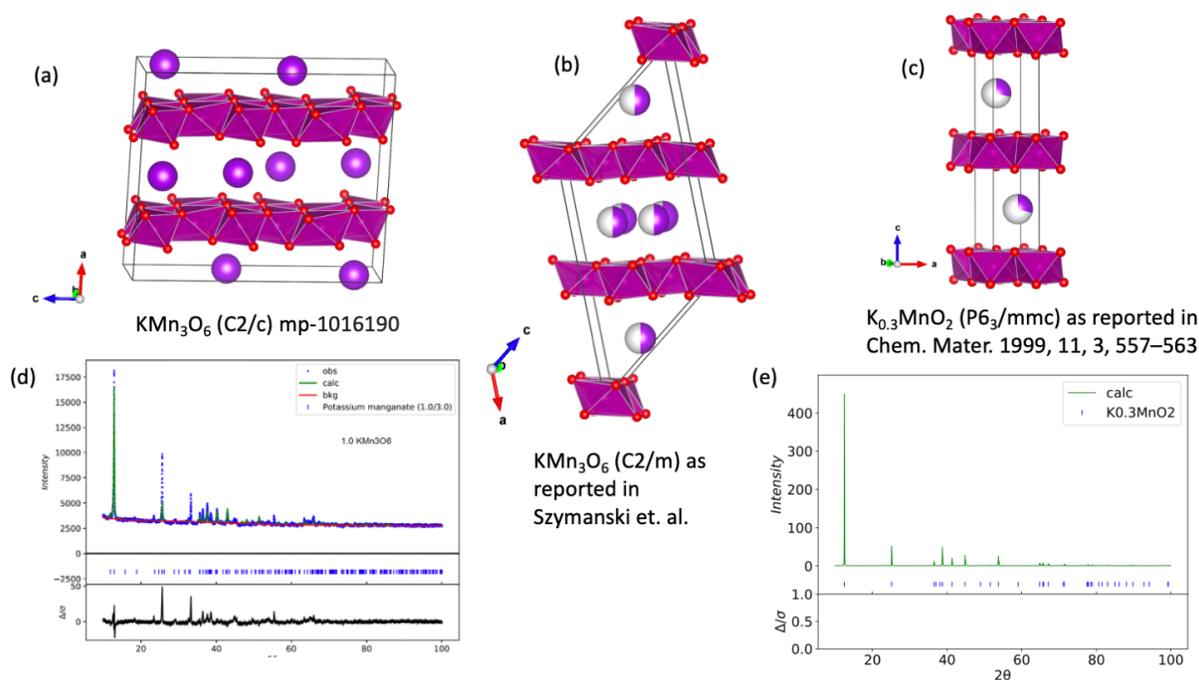


Figure 2: (a) Structure of  $KMn_3O_6$  as predicted by Materials Project. (b) Structure of  $KMn_3O_6$  as provided by Szymanski et al. (c) Structure of  $K_{0.3}MnO_2$  as reported in Kim et al.<sup>22</sup> K is shown as large purple spheres, Mn is in pink octahedra and O is shown in red. (d) PXRD pattern as provided by Szymanski et al. and (e) simulated PXRD pattern for  $K_{0.3}MnO_2$ .

## Pb-Sb pyrochlores

Pyrochlores are structures which have formula  $A_2B_2O_{7-\delta}$ , with A and B denoting two possible cation sites and  $\delta$  the possible oxygen defect.<sup>23</sup> Stoichiometric pyrochlores, where  $\delta = 0$  feature metal ions with one of two combinations of formal charge: A(II)B(V) or A(III)B(IV).

The compounds of interest here all have an A site of Pb(II), and in general lead pyrochlores typically show significant non stoichiometry: A-site vacancies or Pb(IV) defects on the B site compensated by oxygen vacancies, i.e.  $\delta > 0$ . Furthermore, the B site in lead pyrochlores can be occupied by two different cation species; for example if half of the B(V) site is replaced by a M(IV) cation the resulting formula is  $\text{Pb}_4\text{M}_2\text{Sb}_2\text{O}_{13}$  - these compounds are known for  $\text{M(IV)} = \text{Ti, Zr, Hf, Sn}$  and others. Alternatively, one quarter of the B(V) sites can be replaced by M(III) ions, yielding compounds with stoichiometry  $\text{Pb}_4\text{M}_1\text{Sb}_3\text{O}_{13}$ , example M(III) cations that have been incorporated in this way are In, Al, Sc, Cr, Fe, Ga, Rh.<sup>24,25</sup>

The Materials Project predicts the stability of various compounds with the formula  $\text{Pb}_4\text{M}_x\text{Sb}_{4-x}\text{O}_{13}$ , where  $x = 1, 2$  and M is a cation. A-Lab reported the successful synthesis and refinement of five  $\text{Pb}_4\text{M}_x\text{Sb}_{4-x}\text{O}_{13}$  compounds, namely  $\text{M} = \text{Fe, In, } x=1$  ( $\text{FeSb}_3\text{Pb}_4\text{O}_{13}$ , and  $\text{InSb}_3\text{Pb}_4\text{O}_{13}$ ), which are predicted to crystallize in space group  $R\bar{3}m$ , as well as  $\text{M} = \text{Hf, Sn, Zr, } x=2$  ( $\text{Hf}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$ ,  $\text{Sn}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$ , and  $\text{Zr}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$ ), which are predicted to crystallize in space group  $Imm2$ . *N.B.* the naming convention for pyrochlores is that the A site is the larger metal ion of lower charge and appears first in the formula, but to avoid confusion we will use the compound names given in the A-lab paper, which in each case place the B site ion first. It should be pointed out that for each of the five compounds listed above, there exists a reported, disordered version of the material on the ICSD. In fact all of the B sites in question were incorporated into lead antimony pyrochlores by Cascales and coworkers in a series of papers in 1985-86, and each is reported by them as consisting of random B site substitutions, retaining the pyrochlore symmetry  $Fd\bar{3}m$ .<sup>24,25</sup>

Since the predicted structures for these five compounds are just B site ordered versions of the known disordered structures, we again emphasise that evidence of the ordering must be found in the characterisation in order to prove the formation of the ordered phase. We note that the x-ray scattering factors between M and Sb are very similar for  $\text{M} = \text{In, Hf, Sn}$ , meaning that any ordering of these elements in the synthesised materials would be very difficult to detect via XRD.

Table 2: Comparison of the pseudo-cubic lattice parameters predicted in Materials Project with the reported cubic lattice parameters listed in the ICSD for doped pyrochlores.

Compound	Pseudo-cubic lattice parameter from A-lab CIF / Å	Cubic lattice parameter of equivalent disordered phase from ICSD / Å
Pure $\text{Pb}_2\text{Sb}_2\text{O}_7$	-	10.44
$\text{Sn}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$	10.6168	10.5645
$\text{Zr}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$	10.6594	10.6349
$\text{Hf}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$	10.6415	10.6126
$\text{FeSb}_3\text{Pb}_4\text{O}_{13}$	10.4931	10.4803
$\text{InSb}_3\text{Pb}_4\text{O}_{13}$	10.5845	10.5892

The ordered B site cations in the predicted structures lead to a lower symmetry unit cell compared with the cubic pyrochlore of the parent  $\text{Pb}_2\text{Sb}_2\text{O}_{6.5}$ . However, if the atom type is ignored, the atom positions are almost the same as in the cubic pyrochlore. We calculate the pseudo-cubic lattice parameter,  $a_p$ , of the ordered CIFs produced by A-lab, and these are shown in Table 2. Overall these show close agreement with the known phases. The largest difference is the  $\text{Sn}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$  compound which we discuss in detail below. Other compounds have closely matching parameters, and with no evidence of order we conclude they are very likely the known, disordered pyrochlore compounds discovered in the 1980s.

Fig. 3 shows that the calculated diffraction pattern of the known phase  $\text{Pb}_2\text{SnSbO}_{6.5}$  and the predicted  $\text{Sn}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$  are almost completely identical. Although the B site ordering in the predicted phase lowers the symmetry, the intensities of the additional reflections are very weak. As such, alternative analytical steps are required to assert the existence of B site ordered  $\text{Sn}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$ . And most importantly, it needs to be verified that this material is different from  $\text{Pb}_2\text{SnSbO}_{6.5}$ .<sup>26</sup> We find that the Rietveld refinement of the PXRD data from  $\text{Sn}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$ , shown in Fig. 4 can be carried out successfully using the known compounds  $\text{Pb}_2\text{SnSbO}_{6.5}$  and  $\text{SnO}_2$ , both reported on ICSD. Thus using the argument of Occam's razor, we conclude that this is the more likely interpretation of the synthetic products. In general, our attempts at Rietveld refinement of the data from all five of the reported pyrochlore samples, as shown previously in figure 4 and in figures S11, S12, S13, and S15 in the appendix,

indicates that the synthesized phases are likely the known disordered B site pyrochlores crystallizing in the higher symmetry space group  $Fd\bar{3}m$ . Thus, the pyrochlores are all an example of error three, missed disorder, or no evidence of cation order.

Notably, all these materials are related to the famous “Naples Yellow” pigment, which derives from  $\text{Pb}_2\text{Sb}_2\text{O}_7$ .<sup>27</sup> Variants of Naples Yellow, including those with Sn(IV) substitution on the B site, were used by the ancient Egyptians, and have been lost and then rediscovered periodically throughout history, by different ancient civilisations, in the middle ages, at various points in the renaissance, and most recently by the A-lab. Interestingly, it has been debated that  $\text{Pb}_2\text{Sb}_2\text{O}_7$  itself is not stable and that doping (most commonly with Sn, but also other elements) is necessary to stabilize the pigment. For an excellent overview we recommend the following work by Marchetti et. al and the references therein.<sup>27</sup>

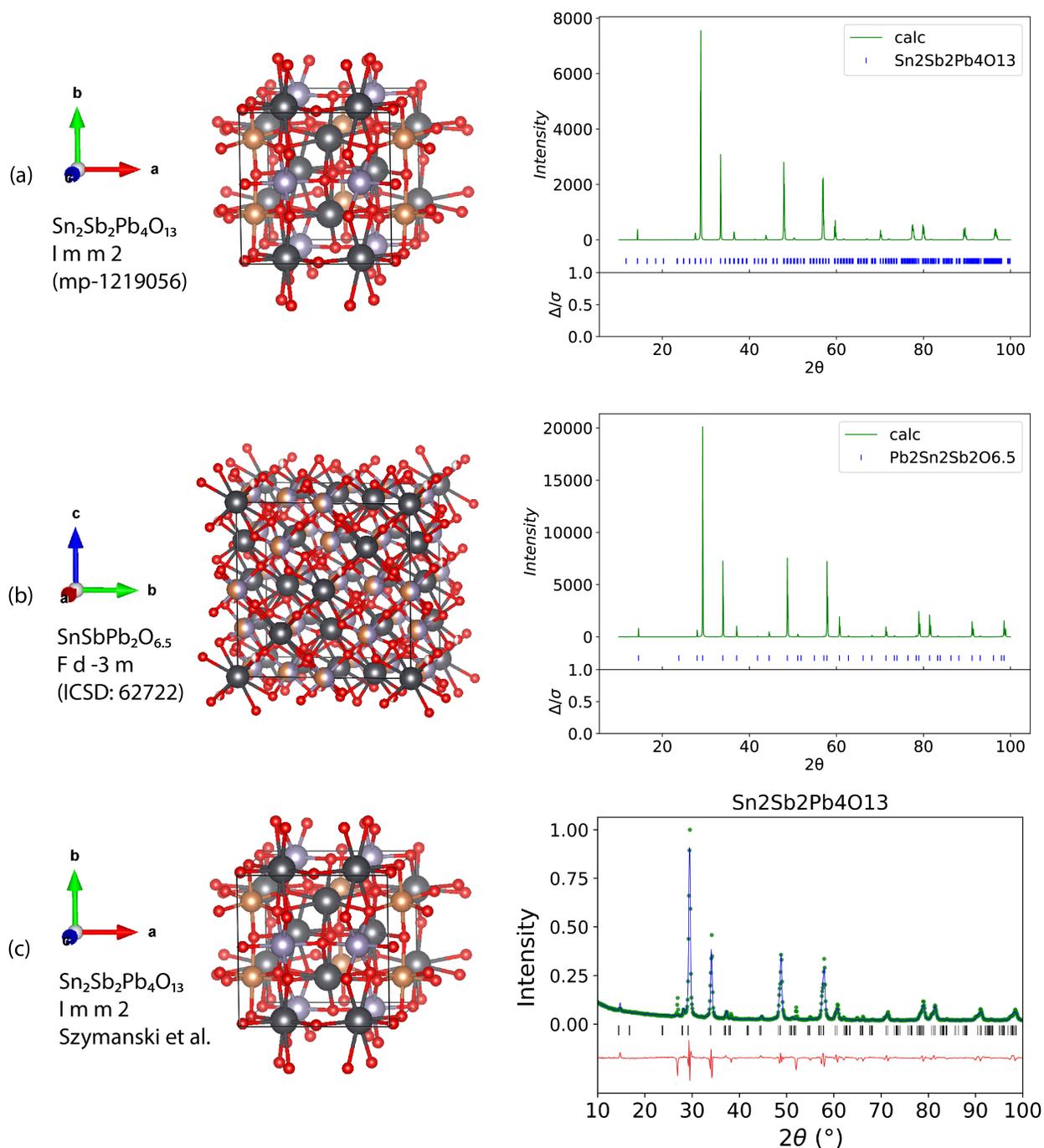


Figure 3: (a) Structure of  $\text{Sn}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$  as predicted by Materials Project (left) and its simulated PXRD pattern (right). (b) Structure of the doped  $\text{SnSbPb}_2\text{O}_{6.5}$  pyrochlore phase as listed in the ICSD (left) and its simulated PXRD pattern (right). (c) Structure of  $\text{Sn}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$  as provided by Szymanski et al. (left) and its reported PXRD pattern (right). Pb in dark grey, Sn in purple, Sb in orange, and O in red.

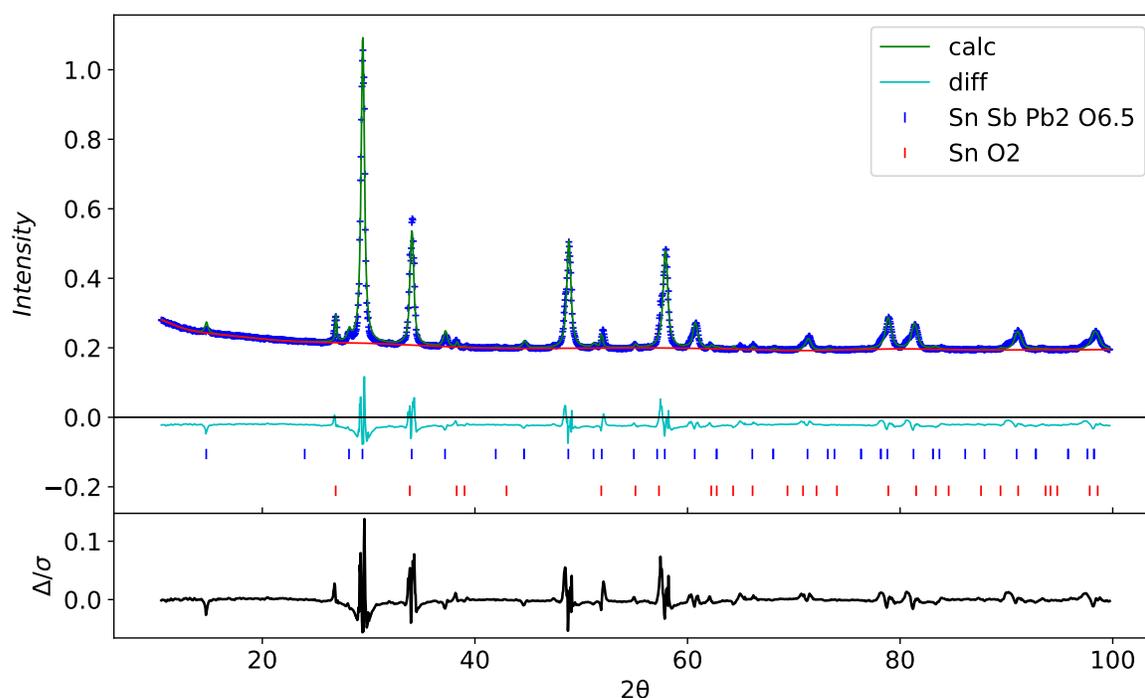


Figure 4: Rietveld refinement of the experimental  $\text{Sn}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$  pattern against the pyrochlore  $\text{SnSbPb}_2\text{O}_{6.5}$  (coll-62722) and  $\text{SnO}_2$  (coll-9163) phases as reported on ICSD. Unlike the predicted  $\text{Sn}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$  phase, which was supposed to crystallize in the orthorhombic space group  $Imm2$ , the pattern can be fit well with the known higher symmetry structure in space group  $Fm\bar{3}m$ .

## Spinel

Spinel, which possess an  $\text{AB}_2\text{O}_4$  stoichiometry, are another common type of oxide materials. A-Lab claimed to have synthesized the following spinel-derived compounds:  $\text{Zn}_2\text{Cr}_3\text{FeO}_8$  ( $R\bar{3}m$ ),  $\text{Zn}_3\text{Ni}_4(\text{SbO}_6)_2$  ( $C2/c$ ), and  $\text{Mn}_4\text{Zn}_3(\text{NiO}_6)_2$  ( $C2/c$ ). However, our refinements indicate that the diffraction patterns can be better or equally well interpreted as known, B site disordered cubic spinels, all of which have been tabulated in the ICSD database.

$\text{NiMn}_2\text{O}_4$  is a cubic spinel with complex cation and magnetic ordering. Guillemet-Fritsch et al. explored Zn insertion into  $\text{NiMn}_2\text{O}_4$ ,<sup>28</sup> including synthesis of a compound with almost exactly the stoichiometry as the A-lab material  $\text{Mn}_4\text{Zn}_3(\text{NiO}_6)_2$ . In the work of Guillemet-Fritsch et al., the A and B sites were found to be compositionally disordered, with some Zn

migration onto the octahedral sites, forming a disordered arrangement with the Mn and Ni ions. However, the prediction from the A-lab is of the A site (tetrahedral site) exclusively occupied by Zn, and the B site (octahedral) with ordered Ni and Mn. In Fig. 5 the PXRD data from the A-lab sample  $\text{Mn}_4\text{Zn}_3(\text{NiO}_6)_2$  are refined against the existing compounds. The predominant phase in this pattern matches well to the disordered cubic spinel phase  $(\text{Zn}_{0.759}\text{Mn}_{0.241})(\text{Mn}_{1.35}\text{Ni}_{0.65})\text{O}_4$  in space group  $Fd\bar{3}m$ , with minority peaks indexing to NiO. The analysis of this compound, again, suffers from error three.

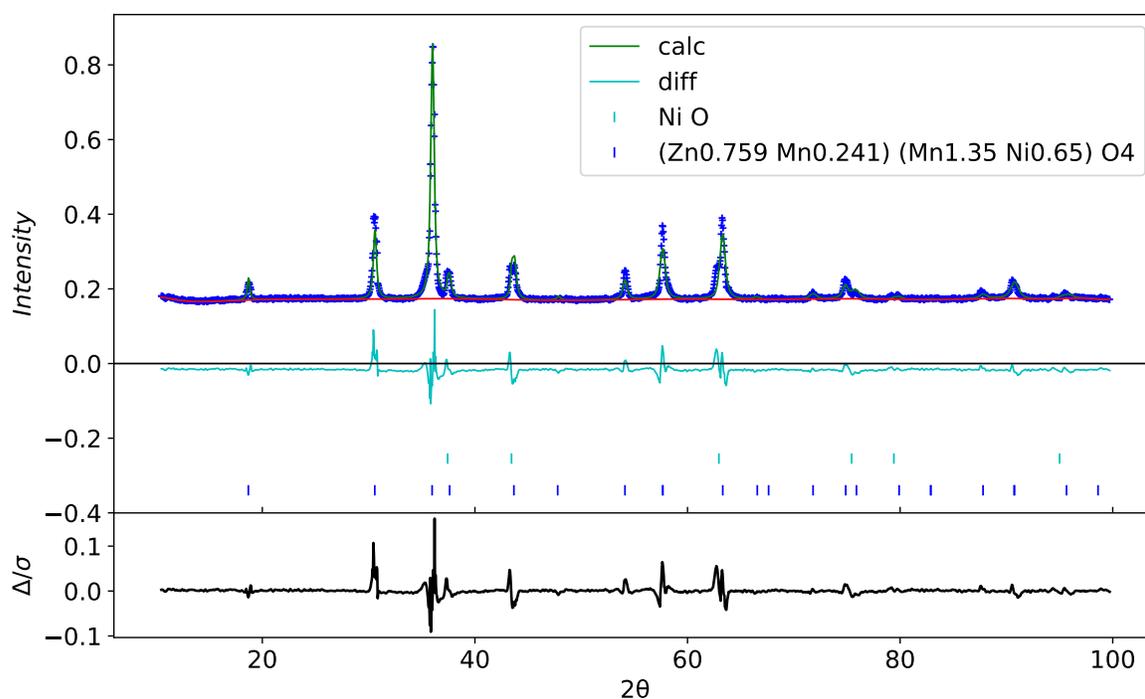


Figure 5: Our Rietveld refinement of  $\text{Mn}_4\text{Zn}_3(\text{NiO}_6)_2$  against the doped spinel phase  $(\text{Zn}_{0.759}\text{Mn}_{0.241})(\text{Mn}_{1.35}\text{Ni}_{0.65})\text{O}_4$  (coll-92223) and NiO (coll-9866) as reported on ICSD.

Another synthesis to consider is the one that targeted  $\text{Zn}_3\text{Ni}_4(\text{SbO}_6)_2$ . Gama et al.<sup>29</sup> studied the inverse spinel  $\text{Zn}_7(\text{SbO}_6)_2$  and Ni substituted variants. As its concentration increases, Ni replaces the Zn on the octahedral site, progressively converting the material to the normal spinel structure, which it attains at the composition  $\text{Zn}_3\text{Ni}_4(\text{SbO}_6)_2$ , i.e. the exact composition predicted in the A-lab paper. Gama et al. found the Ni and Sb to be disordered on the B site. Given the large x-ray scattering factor difference between Sb and Ni,

detection of B site order should be very straightforward by PXRD. The structure obtained from the Materials Project, and described in the paper, mp-1216023, has ordered Ni and Sb ions on the B site. However, in this case, as has been seen previously, the Materials Project structure differs from the structure file provided in the supplementary information. While the symmetry of both the structure predicted in Materials Project and the structure reported in<sup>4</sup> have the same symmetry (space group  $C2/c$ ), in the latter, the Ni and Sb ions are disordered (having fractional occupancies in the CIF). In Fig. 6 we compare (a) the predicted structure from Materials Project (mp-1216023) with its calculated diffraction pattern, (b) the reported ICSD structure from Gama et al.(ICSD-109468) with its calculated diffraction pattern, and (c) and the structure provided in the supplemental information of,<sup>4</sup> with the experimental pattern and A-lab fit. It is clear that the ordering of the B site ions has a very large effect on the calculated PXRD patterns. The ordered pattern has a very strong peak at  $17.9^\circ$ , a feature that is much weaker in the disordered pattern. The experimental pattern clearly matches much better to the disordered version of  $Zn_3Ni_4(SbO_6)_2$ , which has been reported first by Gama et al. in 2003.

Our own refinement is given in Fig. 7. We find it necessary to include impurity phases  $NiSb_2O_6$  (ICSD# 426852), and  $NiO$  (ICSD# 9866) to match all the Bragg peaks. We attempted the fit with Zn-Ni-Sb spinels with different Ni contents from the ICSD. We show the fit with  $Zn_6NiSb_2O_{12}$  (ICSD# 109465), although other Ni contents gave very similar fits and we do not think we can differentiate between them. Therefore, in this case, the ordered B site spinel from the Materials Project is clearly not present, as the strong low angle ordering peak that is predicted to appear at  $17.9^\circ$  is absent from the experimental data. The spinel phase that is present matches well with known Zn-Ni-Sb spinels with disordered B site cations. We conclude in the absence of any further evidence, that this is not a new material.

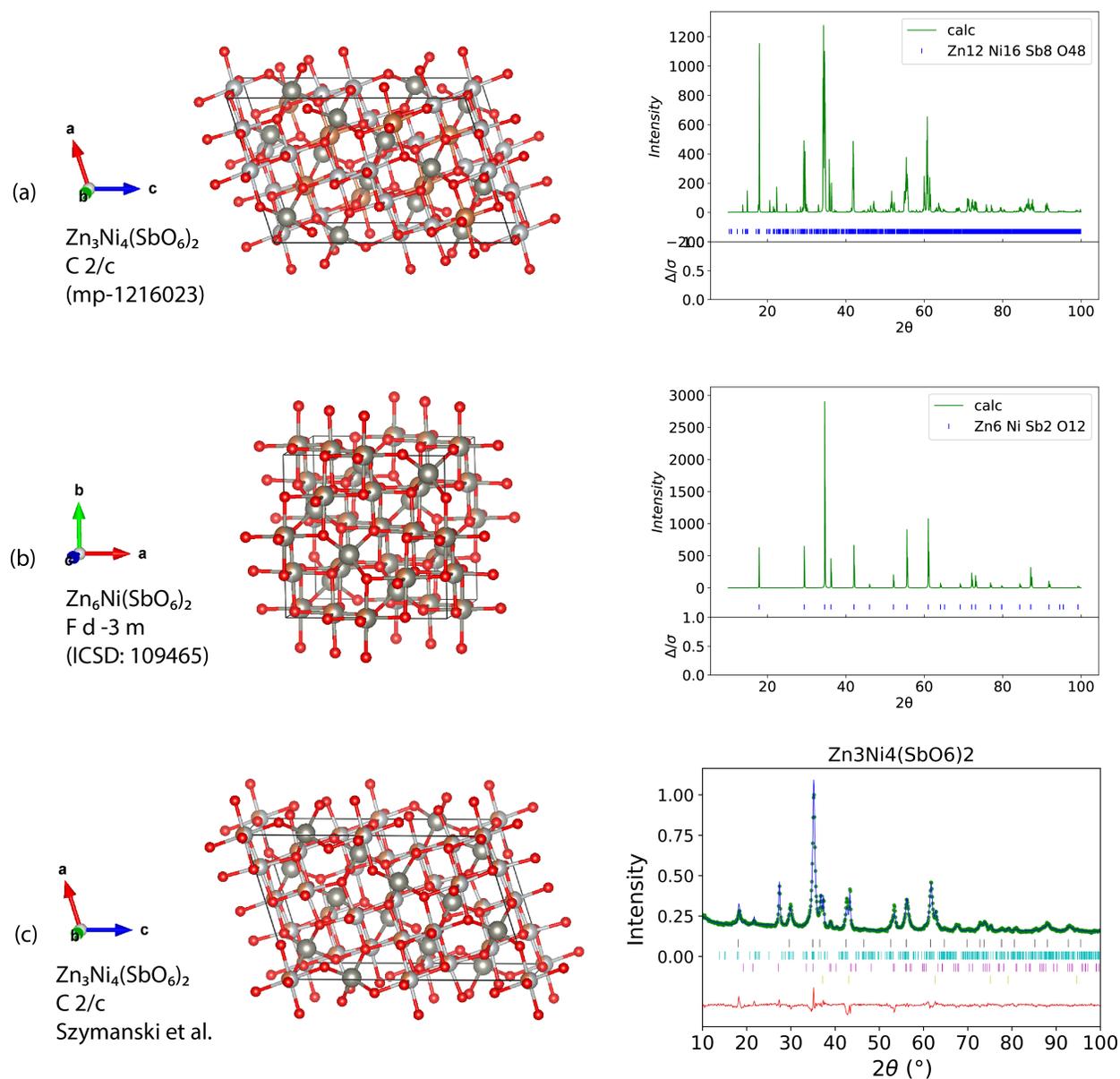


Figure 6: (a) Structure of  $\text{Zn}_3\text{Ni}_4(\text{SbO}_6)_2$  as predicted by Materials Project (left) and its simulated PXR pattern (right). (b) Structure of the doped  $\text{Zn}_6\text{Ni}(\text{SbO}_6)_2$  (coll-109465) phase as listed in the ICSD (left) and its simulated PXR pattern (right). (c) Structure of  $\text{Zn}_3\text{Ni}_4(\text{SbO}_6)_2$  as provided by Szymanski et al. (left) and its reported PXR pattern (right). Ni in white, Zn in grey, Sb in orange, and O in red.

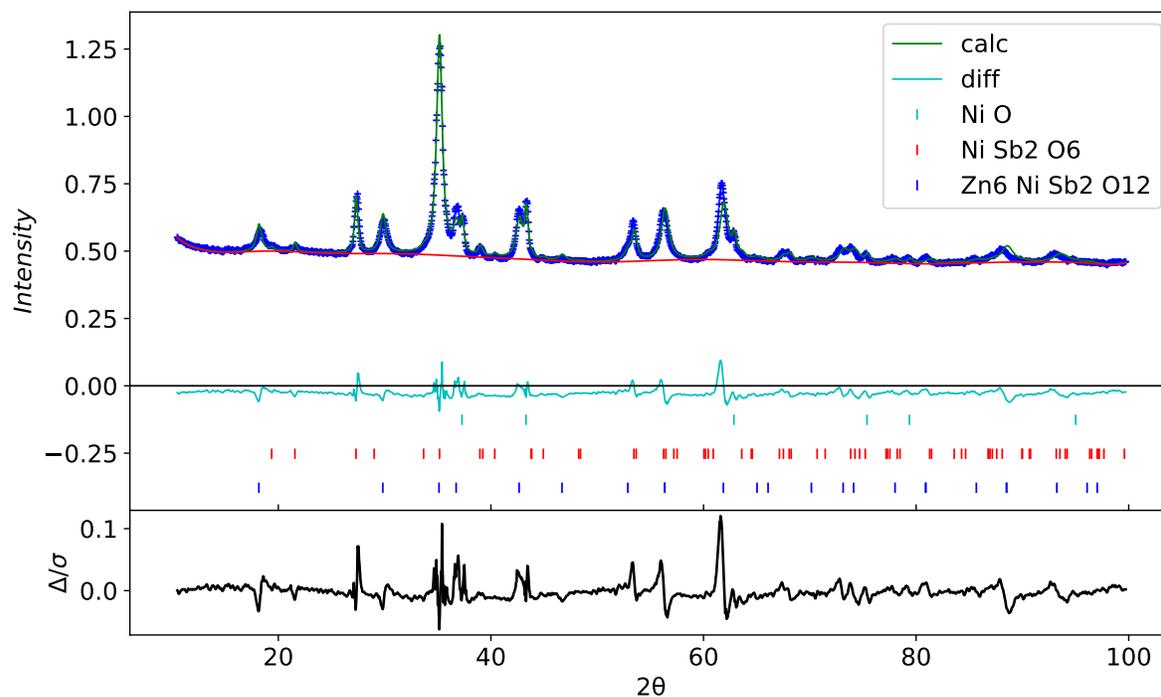


Figure 7: (a) Rietveld refinement of the pattern representing  $\text{Zn}_3\text{Ni}_4(\text{SbO}_6)_2$  against  $\text{Zn}_6\text{NiSb}_2\text{O}_{12}$  (coll-109465),  $\text{NiSb}_2\text{O}_6$  (coll-426852), and  $\text{NiO}$  (coll-9866), as reported on ICSD.

The compound  $\text{Zn}_2\text{Cr}_3\text{FeO}_8$  is predicted by A-lab to have ordered Cr and Fe ions on the B site, but is otherwise identical to the normal spinel structure. A series of solid solutions between  $\text{ZnCr}_2\text{O}_4$  and  $\text{ZnFe}_2\text{O}_4$ , including the exact composition predicted, was reported in 1970,<sup>30</sup> and the B site cations were described as disordered. Due to the small difference in scattering factor between Fe and Cr, there is only minimal difference in PXRD pattern between the predicted, ordered structure and the known, disordered structure. In addition, the reported PXRD pattern of  $\text{Zn}_2\text{Cr}_3\text{FeO}_8$  also contains several unmodelled impurity peaks which likely correspond to binary oxides of the metals. In the absence of any other evidence, we conclude that  $\text{Zn}_2\text{Cr}_3\text{FeO}_8$  is not a new material and is the disordered B site spinel described in 1970. Thus this falls under error number three.

## Perovskites and Ilmenites

Perovskite oxides have the general formula  $ABO_3$ , their characteristic motif is exclusively corner sharing  $BO_6$  octahedra. Ilmenites have the same general formula but the octahedra are edge-sharing and the structure becomes hexagonal or rhombohedral.

A-Lab reported the synthesis of six perovskite- or ilmenite-derived structures:  $Ba_2ZrSnO_6$  ( $Fm\bar{3}m$ , perovskite),  $MgTi_2NiO_6$  ( $R3$ , Ilmenite),  $KBaGdWO_6$  and  $KBaPrWO_6$  ( $F\bar{4}3m$ , double perovskite), as well as  $Ba_6Na_2Ta_2V_2O_{17}$  and  $Ba_6Na_2V_2Sb_2O_{17}$  ( $P6_3/mmc$ , perovskite derivatives). However, our analysis suggest that the powder patterns of these samples more likely correspond to a doped and known perovskite/ilmenite phase mixed with some impurity phases. We will elaborate on this with two detailed examples. We would also like to point to the recent preprint by Yamamoto et al. which provides more insight into the supposed synthesis of  $Ba_6Na_2Ta_2V_2O_{17}$ .<sup>31</sup>

In the case of  $Ba_2SnZrO_6$ , the output of the refinement process, as published in the A-lab paper, contains many significant unmodelled diffraction peaks. Our fit is shown below. We are able to account for all obvious Bragg peaks, and the residuals that remain are characteristic of incorrect peak profiles rather than completely unmodelled diffraction features. Our fit necessitated four phases:  $SnO_2$ ,  $ZrO_2$  (which are both starting materials),  $BaSnO_3$  (present in 44 wt %, a known perovskite phase ICSD 188149) and  $BaSn_{0.5}Zr_{0.5}O_3$  (present in 21 wt %, another known perovskite phase ICSD 43137). Thus in our view the XRD pattern provided is best explained as originating from a mixture of starting materials and known perovskite phases. We note the model used by the authors in their refinement of  $Ba_2SnZrO_6$ , (published as a CIF file) contains disordered Zr and Sn ions on the B site of the perovskite. This contrasts with the Materials Project entry of the stated material, which has ordered Sn and Zr ions, and consequently, a doubling of the lattice parameter. The small difference in x-ray scattering factor between Sn and Zr means that simulated patterns of the ordered  $Ba_2SnZrO_6$  phase show only very small changes compared with the disordered phase. The analysis of  $Ba_2ZrSnO_6$  suffers thus errors one (poor fit), two (inconsistency in predicted and

reported structure) and three (no evidence of cation order).

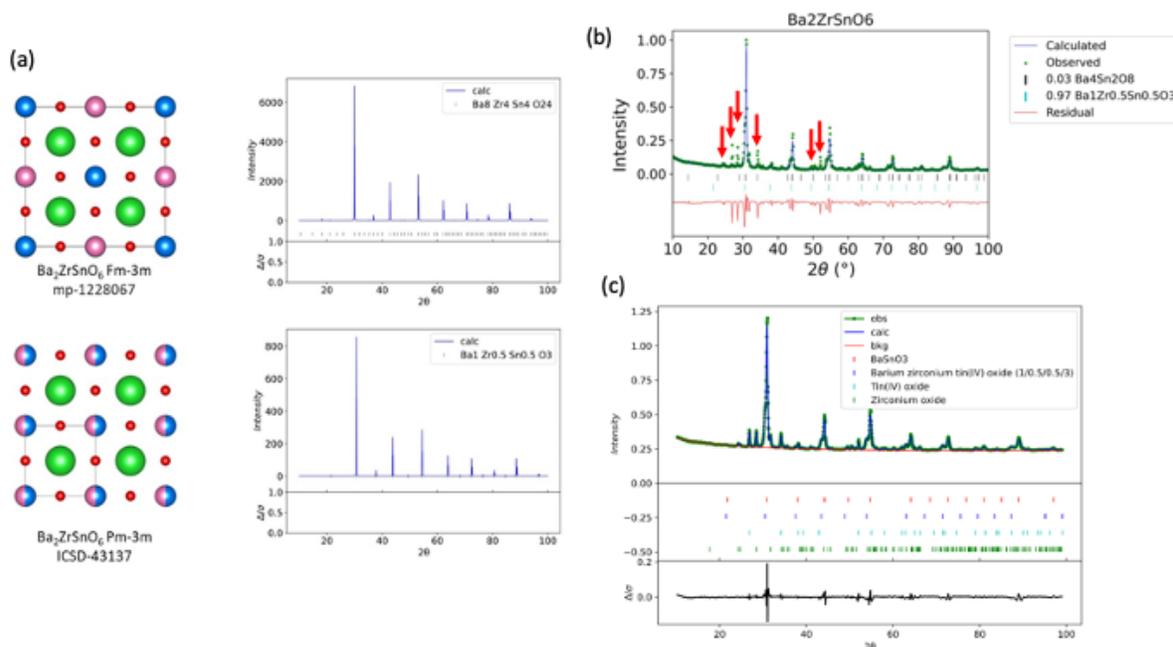


Figure 8: (a) Comparison of the predicted structure of  $Ba_2ZrSnO_6$  with the structure of known  $BaSn_{0.5}Zr_{0.5}O_3$ . The simulated powder patterns of both phases are very similar. (b) refinement as given in,<sup>4</sup> where several peaks are not accounted for (marked by red arrows). (c) Our refinement using  $BaSn_{0.5}Zr_{0.5}O_3$ ,  $BaSnO_3$  and two of the starting materials, yielding a much better fit.

$MgTi_2NiO_6$  with space group  $R\bar{3}H$  is predicted to exist by A-lab as a new compound in a structure that is close to what is known in literature as the  $Ni_3TeO_6$  structure, which can be understood as an ordered ilmenite structure. The ilmenite structure has space group  $R\bar{3}H$ ; the difference in space group between the ilmenite and  $Ni_3TeO_6$  structures is due to the fact that in the ilmenite structure, there is only one crystallographic A and B site, but with the addition of cation ordering, the symmetry changes to  $R\bar{3}H$ , reflecting the fact that there are now two crystallographically distinct A sites and two distinct B sites. The differences are shown in Figure 9. The calculated diffraction patterns of  $MgTi_2NiO_6$  in the ilmenite and  $Ni_3TeO_6$  structures differ only in the intensities of the diffraction peaks – the nature of the ordering does not allow new reflections that are absent in the disordered structure. Similarly to previous examples, for this particular compound, due to the electron

densities of the constituent atoms, the intensity difference is very small between ordered and disordered phase. Therefore in our opinion, to distinguish the ordered from the disordered phase by PXRD will be exceptionally challenging, but if it is to be attempted then very careful measurement of the peak intensities and explicit comparison with the ordered and disordered models need to be made. When we turn to the diffraction pattern presented in the A-lab paper,<sup>4</sup> we find the pattern for  $\text{MgTi}_2\text{NiO}_6$  can be adequately fitted with a model of the known ilmenite phase  $\text{Ni}_{0.5}\text{Mg}_{0.5}\text{TiO}_3$  and a small impurity (2 wt%) of NiO. This again, is an example of error three. It may be that exceptionally careful analysis of the diffraction intensities can provide evidence that the new, ordered phase has been made. But no such evidence is given in the A-lab paper.<sup>4</sup> We conclude that in the absence of evidence of the new phase, the explanation involving known phases is preferred. Again, as for all other examples this does not mean the predicted ordered phase cannot be synthesized. More careful analysis and changes in synthetic recipes can potentially lead to the successful synthesis of the predicted phases in the future.

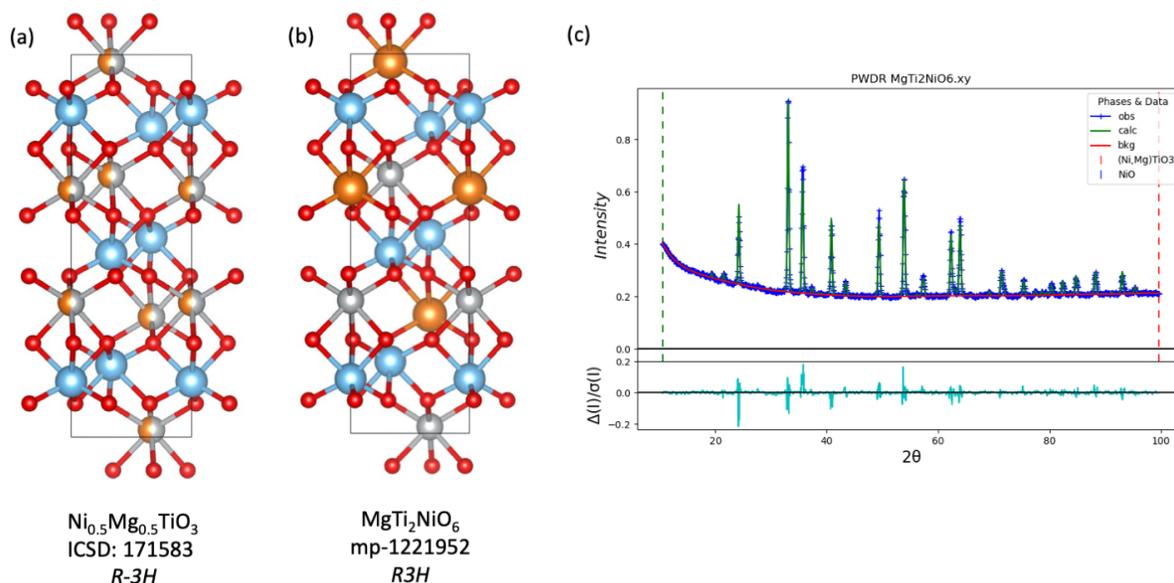


Figure 9: (a) Comparison of the predicted structure of  $\text{MgTi}_2\text{NiO}_6$  with the structure of known  $\text{Ni}_{0.5}\text{Mg}_{0.5}\text{TiO}_3$  (b). (c) Refinement of the corresponding PXRD pattern using  $\text{Ni}_{0.5}\text{Mg}_{0.5}\text{TiO}_3$  yielding a satisfactory fit.

For the double perovskites  $\text{KBaGdWO}_6$  and  $\text{KBaPrWO}_6$ , the fits are so poor as to be meaningless in our view, with clear unfitted Bragg peaks in both cases. Both these structures have order on both the A and B sites (so called double-double-perovskites). A site order in perovskites is much rarer than B site ordering, and is almost always accompanied by oxygen deficiency,<sup>32</sup> a classical example is the cuprate superconductor  $\text{YBa}_2\text{Cu}_3\text{O}_7$ . In the absence of oxygen deficiency, A site order can rarely occur but has to our knowledge not been seen in the symmetry predicted here - typically A sites form ordered layers driven by large charge differences, rather than fcc patterns.<sup>33</sup> Given the novelty of what is being proposed here, clear evidence must be given that the predicted compounds have been formed, with special attention paid to small diffraction features that characterise the ordering. This evidence is not supplied by the poor fits to the PXRD data.

## Phosphates

Phosphates are ionic compounds that contain  $\text{PO}_4^{3-}$  or related anions. They can be highly complex, featuring many crystallographically distinct anions, giving rise to a large unit cell with low symmetry and a complex associated diffraction pattern. The A-lab paper reports 18 ‘new’ phosphate, diphosphate (with  $\text{P}_2\text{O}_7^{4-}$  anions) or metaphosphate ( $\text{PO}_3^-$ ) containing phases. Those are:  $\text{K}_2\text{TiCr}(\text{PO}_4)_3$ ,  $\text{CaFe}_2\text{P}_2\text{O}_9$ ,  $\text{CaCo}(\text{PO}_3)_4$ ,  $\text{CaMn}(\text{PO}_3)_4$ ,  $\text{CaNi}(\text{PO}_3)_4$ ,  $\text{InSb}_3(\text{PO}_4)_6$ ,  $\text{K}_4\text{MgFe}_3(\text{PO}_4)_5$ ,  $\text{K}_4\text{TiSn}_3(\text{PO}_5)_4$ ,  $\text{KNaTi}_2(\text{PO}_5)_2$ ,  $\text{KNaP}_6(\text{PbO}_3)_8$ ,  $\text{MgNi}(\text{PO}_3)_4$ ,  $\text{MgTi}_4(\text{PO}_4)_6$ ,  $\text{Na}_3\text{Ca}_{18}\text{Fe}(\text{PO}_4)_{14}$ ,  $\text{Na}_7\text{Mg}_7\text{Fe}_5(\text{PO}_4)_{12}$ ,  $\text{NaMnFe}(\text{PO}_4)_2$ ,  $\text{Mn}_2\text{VPO}_7$ ,  $\text{Mn}_7(\text{P}_2\text{O}_7)_4$ , and  $\text{MgCuP}_2\text{O}_7$ .

As mentioned previously, one must be careful in refining such complex structures because the abundance of reflections and atomic positions may lead to meaningless fits that have too many parameters. As such, a pure sample, high quality PXRD data, and other identification methods are imperative to assert the synthesis of a new phase, if the structure is highly complex and of low symmetry. Errors one (poor fit) and three (no evidence of cation order) are very common in the analysis of the phosphates.

The compound  $\text{K}_2\text{TiCr}(\text{PO}_4)_3$  was predicted to exist as a new cubic phase in the space group  $P2_13$ . Fig. 10(d) shows our refinement of the provided PXRD pattern, using known cubic  $\text{K}_2\text{Ti}_2(\text{PO}_4)_3$  ( $P2_13$ ; ICSD # 202888) and  $\text{Cr}_2\text{O}_3$ , a common impurity in high temperature synthesis of oxides containing chromium.<sup>34</sup> The refinement provided in the A-lab paper had several unfitted peaks, which all correspond to the  $\text{Cr}_2\text{O}_3$  impurity phase as marked by red arrows in Fig. 10(c). The example of  $\text{K}_2\text{TiCr}(\text{PO}_4)_3$  shows that there are serious issues with the supposed synthesis of the phosphates, in fact we could index and preliminarily match all 18 PXRD patterns to materials that are reported in the ICSD (see Table S3 in the supplemental information). We consider it to be the responsibility of the authors of the A-lab paper<sup>4</sup> to unambiguously prove the synthesis of the target materials in all cases and will refrain from providing alternative refinements of all 43 materials in this comment. We will however discuss each compound and possible alternatives briefly below.

$\text{KNaP}_6(\text{PbO}_3)_8$  was predicted to be in space group  $P3$  with ordered vacancies that leave tunnels in the structure. The structure provided in the supplemental information of the A-lab paper<sup>4</sup> has a higher symmetry and the space group  $P6_3/m$ , although the authors incorrectly claim they synthesized the target compound in space group  $P3$ . The structure provided in the supplemental information not only has a higher symmetry, it also places K, Na and Pb on the same crystallographic position. We found that the PXRD pattern of the synthetic product could also be indexed with a combination of the known materials  $(\text{Na,K})\text{Pb}_4(\text{PO}_4)_3$  and  $\text{Pb}_8\text{O}_5(\text{PO}_4)_2$ , which provides an alternative interpretation of the data.

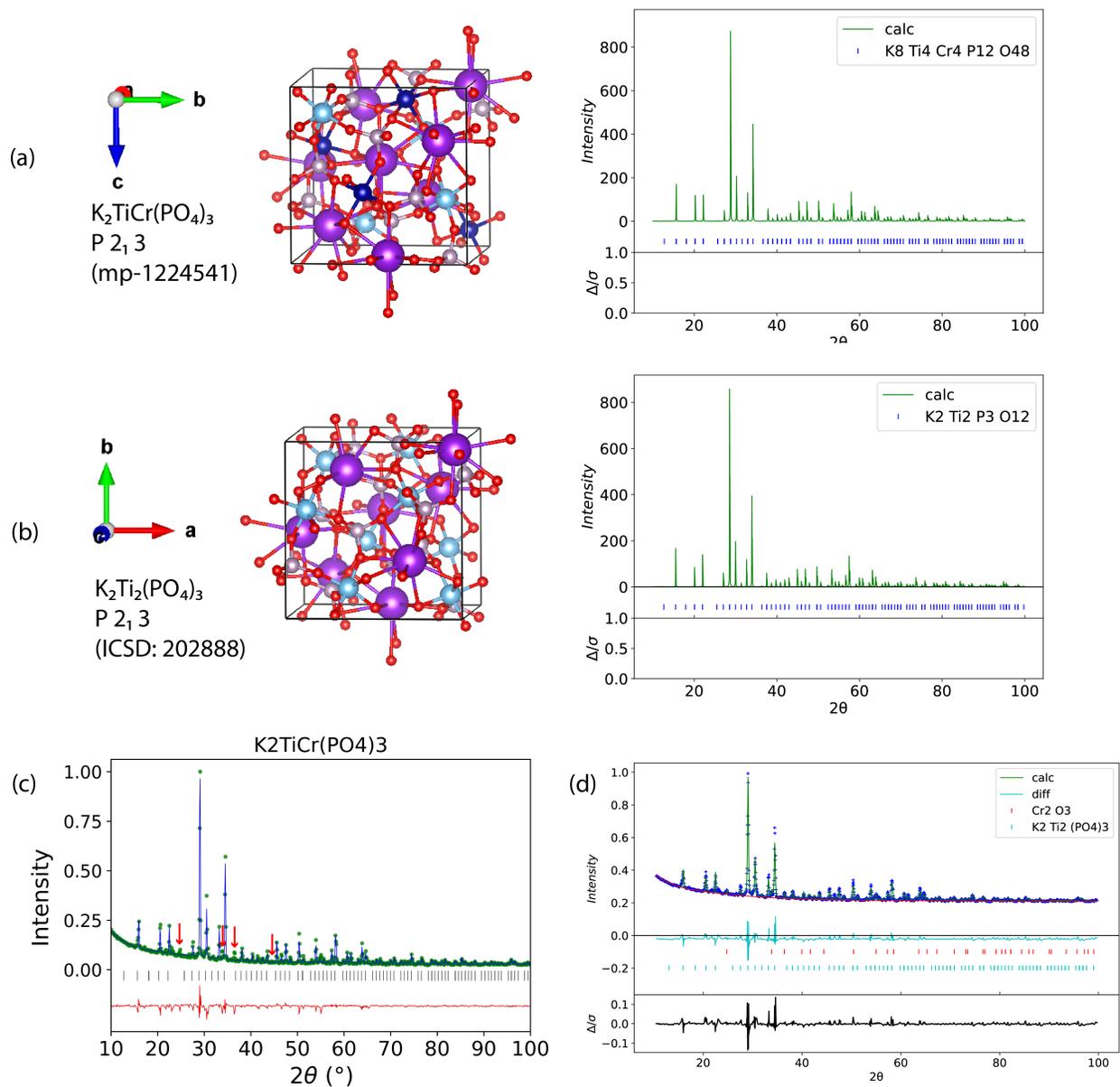


Figure 10: (a) Structure of  $K_2TiCr(PO_4)_3$  as predicted by Materials Project (left) and its simulated PXR pattern (right). (b) Structure of  $K_2Ti_2(PO_4)_3$  phase as listed in the ICSD (left) and its simulated PXR pattern (right). (c) The reported XRD pattern of  $K_2TiCr(PO_4)_3$  as provided by Szymanski et al. with the unmatched peaks marked by arrows in red. (d) Our refinement of the experimental  $K_2TiCr(PO_4)_3$  pattern against the  $K_2Ti_2(PO_4)_3$  phase (coll-202888) and  $Cr_2O_3$  (coll-626479) as reported on ICSD. The unmatched peaks in the refinement provided in the A-lab paper are matched to the  $Cr_2O_3$  phase. K in purple, Ti in sky blue, Cr in navy blue, P in light purple, and O in red.

There are four predicted compounds in the A-lab dataset based on the  $Ni_2(PO_3)_4$  struc-

ture, a kind of tetrametaphosphate with two inequivalent metal sites, both octahedrally coordinated but with slightly differently sized coordination environments. This  $M_2(PO_3)_4$  structure type is known to form with a variety of M(II) cations, namely  $M = Mg, Mn, Fe, Co, Ni, Cu, Zn$ . These previously reported compounds are listed in the ICSD. The fact that there are two distinct metal sites offers a possibility that two metals could order over these sites. A few of these bimetallic tetrametaphosphates  $((M,M')_2(PO_3)_4)$  have been synthesised previously, namely  $(M,M') = (Ni,Zn), (Ni,Co)$  and  $(Mg, Mn)$ . Nord investigated the ordering for bimetallic tetrametaphosphates with  $(M,M') = (Ni,Zn), (Ni,Co)$ , using neutron diffraction, as XRD has only small scattering factor differences between these metals.<sup>35</sup> It was found that there is a slight preference for the Ni(II) ion to occupy the smaller octahedral site due to its lower ionic radius, and thus some ordering of the cations is observed. In contrast, in a recent study on  $MgMn(PO_3)_4$  using XRD, no cation ordering was observed.<sup>36</sup> Experimental data is presented by the A-lab for  $CaCo(PO_3)_4, CaMn(PO_3)_4$  and  $CaNi(PO_3)_4$ . It should be noted that unlike for the bimetallic tetrametaphosphates discussed above, which could be thought of as solid solutions between two known  $M_2(PO_3)_4$  compounds, the pure  $Ca_2(PO_3)_4$  compound in this structure is not known, perhaps because Ca(II) is much larger than any of the other M(II) ions listed above that can form this structure. Despite this, it might be possible that Ca(II) could enter into the  $(M,M')_2(PO_3)_4$  structure, and could order in the way seen in reported compounds. However, it is also possible that in fact no Ca has entered the tetrametaphosphate phase, as the PXRD pattern of the bimetallic phases are almost identical to that of the single metal analogue. For example, the proposed  $CaNi(PO_3)_4$  simulated PXRD is almost identical to that of the known compound  $Ni_2(PO_3)_4$ . Without compositional information, careful measures of intensity of the PXRD peaks, or use of alternative diffraction techniques such as neutron diffraction, there is no way to confirm whether the Ca containing compounds have been made. The same is true for the last of these compounds,  $MgNi(PO_3)_4$  - although this is a combination of two metals known to form tetrametaphosphates, there is no way to know from the data presented that the material

produced contains both metal ions, or whether they are ordered as predicted. The fit for this particular sample is also much poorer than the others discussed in this paragraph.

The compound  $\text{NaMnFe}(\text{PO}_4)_2$  is predicted as a kind of olivine structure, but the diffraction pattern is not well fitted by this model, and resembles more closely an Alluaudite phase, an orthophosphate structure, of which many compounds are known containing Na, Mn, Fe metals in different proportions, which are included in the ICSD.<sup>37</sup> We consider this the more likely identity of the material.

The compound  $\text{Mn}_2\text{VPO}_7$  is the Thortveitite structure, which is known for both  $\text{Mn}_2\text{P}_2\text{O}_7$  and  $\text{Mn}_2\text{V}_2\text{O}_7$ . It is reasonable to suggest that a solid solution might be formed between these two phases, representing a phosphate-vanadate compound that does not seem to be in the literature. However, as in many cases here, to differentiate between the predicted phase with ordered vanadate and phosphate ions, and either of the known compounds  $\text{Mn}_2\text{P}_2\text{O}_7$  or  $\text{Mn}_2\text{V}_2\text{O}_7$  by PXRD requires careful measurement of the peak intensities, which change only a small amount as V is reasonably close to P in electron density. This analysis was not done, and the peak fit shows clear deviations from the observed intensities. Therefore no evidence for the formation of the mixed anion phase, nor for ordering of that phase, has been provided.

The predicted compound  $\text{MgTi}_4(\text{PO}_4)_6$  is a known compound, reported by Barth et al. in 1993.<sup>38</sup> The proposed structure has ordered Mg ion vacancies. If the Mg ion vacancies were disordered, then the small peak seen experimentally at around  $17^\circ$  would be absent. This peak was also observed by Barth et al, who discussed the interpretation of this peak as being due to partial ordering of the Mg vacancies. Therefore, neither the compound nor the reported (partial) ordering is new. In addition, a second structure is known to form from this composition, reported in 2008 in a different structure.<sup>39</sup>

The predicted compound  $\text{KNaTi}_2(\text{PO}_5)_2$  is a cation ordered version of a series of titanyl phosphates that have been studied by various groups, for example Norberg et al. reported one such disordered version with a very similar stoichiometry to that predicted here in 2003.<sup>40</sup>

The predicted ordering causes only minimal changes to the PXRD peak intensities, and while it may be possible to discern it with careful measurements, and comparison between models, these were not done, so there is no evidence of the ordered phase. It is likely the previously reported disordered phase was made.

The predicted compound  $\text{MgCuP}_2\text{O}_7$  is based on mp-1041741. However the CIF file provided in the Supplementary Information has disordered cations while the Materials Project CIF has ordered Mg and Cu cations. The simulated diffraction pattern fits much better for the disordered version.

To complete this section, we feel the fits to the models for the predicted compounds  $\text{Na}_3\text{Ca}_{18}\text{Fe}(\text{PO}_4)_{14}$ ,  $\text{Na}_7\text{Mg}_7\text{Fe}_5(\text{PO}_4)_{12}$ ,  $\text{Mn}_7(\text{P}_2\text{O}_7)_4$ , and  $\text{K}_4\text{TiSn}_3(\text{PO}_5)_4$  are so poor that they cannot provide any evidence of formation of a new material, in such a complex composition space with a large number of possible candidate phases. We will not explore these particular compounds further.

## Other Materials

Six of the newly reported materials do not fit in the categories we established above and we will discuss them here separately. These are  $\text{MgV}_4\text{Cu}_3\text{O}_{14}$  (*P1*),  $\text{CaGd}_2\text{Zr}(\text{GaO}_3)_4$  (*P1*),  $\text{Ba}_9\text{Ca}_3\text{La}_4(\text{Fe}_4\text{O}_{15})_2$  (*P1*),  $\text{KNa}_2\text{Ga}_3(\text{SiO}_4)_3$  (*P2*<sub>1</sub>/*c*),  $\text{KPr}_9(\text{Si}_3\text{O}_{13})_2$  (*P3*), and  $\text{NaCaMgFe}(\text{SiO}_3)_4$  (*C2*). With the exception of  $\text{KPr}_9(\text{Si}_3\text{O}_{13})_2$ , these materials have in common that they resemble highly complex, low symmetry structures. Thus a similar argument as made above for phosphates is valid here: In such low symmetry structures, a large amount of fitting parameters, makes refinement extremely challenging and a lot of care needs to be taken to ensure the low symmetry is real. The data provided does not satisfactorily prove the existence of these phases, and for all low symmetry phases above, we found alternative matches to the PXRD data, which could explain the synthetic product to be a combination of known materials. For example,  $\text{CaGd}_2\text{Zr}(\text{GaO}_3)_4$  is likely a type of garnet, e.g. cubic  $\text{Ca}_{0.95}\text{Zr}_{0.95}\text{Gd}_{2.05}\text{Ga}_{4.05}\text{O}_{12}$  (ICSD #202850, space group *Ia* $\bar{3}$ d) matches the main peaks

of the pattern well.  $\text{NaCaMgFe}(\text{SiO}_3)_4$  is predicted based on mp-1221075. However, the Materials Project entry has ordered cations while the CIF provided in the Supplementary information has all four metals completely disordered. This material is in fact the Pyroxene structure, and this structure with many different ratios of Na-Ca-Mg-Fe are listed in the ICSD (e.g. ICSD #417169). Lenaz et al. described one study into these compounds, which have two crystallographic cation sites, one of which has disordered Ca and Na, and the other has disordered Mg and Fe.<sup>41</sup> The experimental pattern from the A-lab matches well with the known compounds.

$\text{KPr}_9(\text{Si}_3\text{O}_{13})_2$ , which is predicted to have higher symmetry ( $R\bar{3}$ ), again possesses disorder in the CIF file provided in the A-lab paper<sup>4</sup> (between K and Pr), which was not predicted in the original structure. It thus follows a similar theme to many compounds that we have already discussed in detail.  $\text{MgV}_4\text{Cu}_3\text{O}_{14}$  is the same composition as an existing compound in ICSD,  $\text{Cu}_{1.5}\text{Mg}_{0.5}\text{V}_2\text{O}_7$  (ICSD 69731). The predicted compound has ordered Mg and Cu ions, and is based on the  $\text{Cu}_2\text{P}_2\text{O}_7$  structure, whereas the reported version is the same but with Mg and Cu ions disordered. The main difference in the PXRD pattern due to ordering is the presence of a new peak at  $9.3^\circ$ , that is absent in the disordered pattern. Sadly, the A-lab data does not go below  $10^\circ$ , so this confirmatory peak was not measured. There is no evidence from the PXRD data that the ordered compound has been made over the disordered.

## Successfully synthesized materials

In our view, three materials have been successfully synthesized as predicted. All of them, however, have been reported in the literature before. They are  $\text{MnAgO}_2$ ,  $\text{Y}_3\text{In}_2\text{Ga}_3\text{O}_{12}$  and  $\text{CaFe}_2(\text{PO}_4)_2\text{O}$ , which have been reported in the following references respectively.<sup>42-44</sup> Of those  $\text{CaFe}_2(\text{PO}_4)_2\text{O}$  seems to have been convincingly synthesized based on the provided PXRD data, whereas the other two's PXRD patterns are fitted so poorly that it is difficult to state whether the materials indeed have been synthesized. But as those are known

compounds, it is easier to believe based on indexing that those phases can be present in the PXRD, perhaps in combination with impurity phases. In any case, the compounds in question were reported relatively recently, between 2021 and 2023. In fact, the authors of the Google DeepMind paper<sup>3</sup> clarified that they took snapshots of the ICSD in 2021 and thus did not include materials discovered since in their training set. They rightfully view it as a success that materials they predicted based on a 2021 snapshot were since discovered.

## Overview

Above we lined out issues with the ‘new’ materials synthesized by A-lab. Of the 36 compounds classed as ‘successes’ and 7 classified as ‘partial successes’ by A-lab (43 total), we find significant issues with 42 of them (the exception being  $\text{CaFe}_2(\text{PO}_4)_2\text{O}$ ). Thus our view of the success rate is significantly different to the claimed 78%. As discussed above, we could agree that three materials were correctly synthesized, which includes two other known compounds,  $\text{MnAgO}_2$  and  $\text{Y}_3\text{In}_2\text{Ga}_3\text{O}_{12}$ . In this case, the success rate would be 3/58, or 5%, which is far away from the claims in the paper. A more strict interpretation of materials discovery (but one usually applied to traditional laboratories claiming a new discovery), that a new material must not have been published at all previously, lead to the conclusion that none of the materials in the A-lab paper would be counted as discoveries.

We found systematic issues in the analysis of the PXRD patterns, which show that, in this case, AI has failed to correctly derive conclusions from the data. One common error we found might also point to a general issue with the material prediction part of the work, namely, they can often be derived from known compounds in which cation order breaks symmetry. As DFT cannot model compositionally disordered atoms easily, there might be an underlying error in the way those materials are predicted, both in the Google AI paper and the Materials Project. Still, more work might be needed to verify this concern. Researchers using this data should keep this in mind, and this point is expanded upon below.

We assert that A-lab has not successfully synthesized the vast majority of the claimed new compounds. Since we raised issues in the paper shortly after publication, the Ceder group has conceded that A-lab does not live up to human standards, but still claim that “the system offers a rapid way to prove that a substance can be made — before human chemists take over to improve the synthesis and study the material in more detail.”<sup>45</sup> We hope that our comment made it clear that this statement is not justified - the A-lab paper does not provide proof that the new materials can be made. The data can in at least 40 out of 43 cases be interpreted with the opposite conclusion, which is that in these cases, the predicted materials have not been successfully synthesized. Major improvements to the data analysis and the addition of careful compositional characterization are necessary to draw the alleged conclusions.

## Outlook

Here we discuss some perspectives on computational design of new materials, and automated labs for inorganic synthesis, from the point of view of experimental solid state chemists.

Accurate, unsupervised AI whole pattern refinement of diffraction patterns is an important goal for the development of closed loop automated synthetic laboratories. While the ability to quantitatively measure the goodness of fit of a model to the data appears very attractive from the point of view of automation, it is vital to realise that no value of  $R_{wp}$  or  $\chi^2$  alone can ever be a sufficient to conclude that a model is even approximately correct. B. Toby stated that “the most important way to determine the quality of a Rietveld fit is by viewing the observed and calculated patterns graphically and to ensure that the model is chemically plausible.”<sup>46</sup> What role does this ‘human intuition’ in assessing the quality of a fit play, and how can this be replicated by ML models? To begin, we can consider why goodness of fit statistics alone are insufficient. All goodness of fit is not equal: in a Rietveld refinement, completely unfitted diffraction features may increase the  $\chi^2$  by the same amount

as slightly incorrect peak shapes, but normally the former are a much more serious concern for the analyst, as they represent at best a missing component of the sample phase composition, or at worst show that the entire model is wrong, for example, that the real material has a different symmetry to the one modelled.

To understand what more is needed beyond statistical goodness of fit, it is worth recalling that, at its heart, science is the making and testing of hypotheses. In a Rietveld refinement, the immediate hypothesis being tested is whether the model loaded in the software, together with the profile and background functions, mathematically fits the experimental data. But for practical purposes that is never the entire hypothesis under examination by the scientist; they are concerned with broader questions. As we have seen in the discussion above, if the wider claim is that a new material has been discovered, it is not enough to show a good match between the new material and the data — the fit must be better than that obtained for known materials that are likely to be present.

To give a more specific example, if the hypothesis is that a material has ordered cations, then the experimental diffraction pattern should be considered against candidate models with both ordered and disordered cations. In the perovskite structure, the A and B cation sites are distinguished, and usually occupied by large, low charge ions on the A site, and small, higher charge ions on the B site. Differences in size and charge mean that A and B cations seldom mix to any appreciable extent, and rarely would it be necessary to test a model of say, the perovskite  $\text{SrTiO}_3$  with Sr(II) and Ti(IV) disordered. But if two cation types are present on the B site, with similar size and charge, these may well mix, or they may order, and in this case consideration of both ordered and disordered models becomes essential. A well known example from the literature is  $\text{Sr}_2\text{FeMoO}_6$ , where the Fe and Mo B site cations can show different degrees of disorder.<sup>47</sup> Any diffraction peaks that arise from the ordered but not disordered structure, or any peaks that change intensity appreciably between the two models, are clearly of prime importance. The absence of an expected ordering peak, even if that peak is small compared with the other diffraction features, and its absence

perhaps makes only a small difference to the numerical goodness of fit, can be fatal for the hypothesis of cation ordering. The answer to the entire research question may depend on the presence or absence of relatively small diffraction features. In other situations, a peak of exactly the same size, that belongs to a minor impurity phase, may have almost no relevance to the overall research question, beyond suggesting that a small adjustment of the synthetic procedure is needed.

Thus the statistics alone can never capture the full meaning of the fit. In judging the quality of a Rietveld fit, an expert human practitioner will not only look at the statistics and judge the correctness of the chemistry, but also consider what possible alternative models need to be compared against, and, ultimately, how the PXRD evidence helps address the wider research question. This is the capability that unsupervised AI Rietveld Refinement systems must possess to avoid incorrect interpretations, and to truly operate without human intervention.

Several attempts at autonomous interpretation of PXRD data have been made. Mayo et al. match experimental patterns of organic polymorphs to a database of calculated structures.<sup>48</sup> Lunt et al extended this methodology to identify polymorphs of organic materials crystallised and analysed in a robotic lab.<sup>49</sup> Salgado et al. used machine learning to classify crystal system and space group from experimental PXRD patterns.<sup>50</sup> The three aforementioned studies are important advances, and a large part of their value is in clearly outlining the limitations of the methodology. Each has a far more modest goal than unsupervised Rietveld Refinement of unknown compounds.

We have shown throughout our analysis here that the problem of compositional disorder is not well addressed by the methods used in the A-lab paper, with many predicted ordered compounds likely to be known, disordered analogues. It should be considered how fundamental this problem is to the materials prediction field. As discussed earlier, compositional disorder is well accommodated by the equations of crystallography: fractional occupancies can be entered in the structure factor equation. However, fractional occupancy presents

difficulties for many computational methods. This issue has long been recognised. An early approach to model disordered materials is the virtual crystal approximation, (VCA), where virtual atoms, are placed on sites of fractional occupancy, with properties intermediate between the real atoms that share the site.<sup>51</sup> While the VCA has been used to calculate some compositionally disordered oxides, it is generally recognised as having important limitations.<sup>52</sup> The bonding properties of compositionally disordered materials often cannot be successfully modelled by averaged virtual atoms, and this can lead to very large divergence between VCA and experiment.<sup>53</sup> An alternative to VCA, the Coherent Potential Approximation (CPA), was introduced in the 1960s, and uses an effective medium to model the average composition of the disordered material, although this approach is computationally expensive, is incompatible with many implementations of DFT, and for some systems cannot give quantitative results.<sup>54</sup> Both VCM and CPA in different ways look to take averages to represent disorder. Alternatively, disorder may be represented only using whole atom occupancies, avoiding fractional occupancies entirely; there are several methods currently in use of this type. The cluster expansion (CE) model began with work to use the Ising model of magnetism to describe compositional order/disorder in the 1950s,<sup>55</sup> and saw major advances from the 1980s<sup>56</sup> onwards. CE considers finite size clusters and computes the properties of the material as a combination of these. Zunger et al. introduced the Special Quasi-random Structure (SQS) as a way of approximating random distributions of atoms within a finite supercell.<sup>57</sup> Grau-Crespo et al used the concept of calculating all possible configurations in a given cell size that represent the total composition of a disordered material.<sup>58</sup> This can be effective for low doping levels, but computation costs increase as the composition of the disordered site approaches 0.5. Very recently, an approach to model the energetics of compositional disorder using machine learning, instead of DFT, has been published.<sup>59</sup> This very brief survey is provided to show that the issues presented by modelling disordered materials are not new. To our knowledge, the Materials Project does not use any of the approaches we mention above to model disorder. While this may provide significant advantages in economy

of computation, it admits the possibility that any predicted order will be artificial and not be seen in experiment. That two-thirds of the A-lab 'successful' syntheses are likely unrecognised disordered compounds that have been incorrectly modeled as ordered, seems to imply that this limitation of the Materials Project calculation methodology has not been fully appreciated. Indeed in the Google Deepmind GNoME paper, no mention of compositional disorder is made at all.

To conclude, we give some short recommendations that we believe emerge from this episode, for those working in autonomous labs, inorganic materials prediction, and related fields.

1. When claiming new materials have been made or predicted, one must state in which way they are new. This was clearly done by the A lab paper (absence from ICSD), but some may take issue with this definition.
2. When predicting new, non-molecular inorganic materials computationally, the possibility of compositional disorder should be considered explicitly.
3. When analysing characterization data, statistical models of goodness of fit cannot be relied upon alone as a measure of success.
4. Compositional measurements of new materials are as important as structural ones.
5. Avoid inverse Occam's Razor:<sup>60</sup> positive hypothesis testing should not be used in materials discovery, one must also assess the possibility that in fact a known material has been made. Evidence for the novelty of a material must be presented in the context of the answer to point 1. E.g. if the novelty rests in the cation order, evidence must be presented for that.

Finally, it seems clear that robotic labs and AI both will play an important part in the future of materials discovery and solid state chemistry. At this current time, we see two important bottlenecks hindering high-throughput materials discovery. The first is issues

with the prediction of new materials, where tensions between high-throughput and high-quality calculations remain unresolved. The other bottleneck is sample analysis. To predict and produce many samples autonomously is impressive, but if the rate determining step is human-operated analysis, the AI enabled robotic lab may not move faster than a traditional one. Important steps towards unsupervised analysis have been made, but in our view, truly autonomous materials analysis remains a target for future work, or a better developed human-machine interface might drastically help to speed up the process. We hope this comment outlines some of the pitfalls and will help to strengthen this aspect in the future.

## Acknowledgement

LMS was supported by the NSF through NSF grant OAC-2118310 and NSF-MRSEC through the Princeton Center for Complex Materials DMR-2011750, as well as by the Gordon and Betty Moore Foundation's EPIQS initiative through Grant No. GBMF9064 and the David and Lucille Packard foundation. SBL is supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. (DGE-2039656). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. YL was supported by the China Scholarship Council.

## References

- (1) Levin, I. Inorganic Crystal Structure Database(ICSD). 2020; <http://icsd.fiz-karlsruhe.de>.
- (2) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project:

- A Materials Genome Approach to Accelerating Materials Innovation. *APL Materials* **2013**, *1*, 011002.
- (3) Merchant, A.; Batzner, S.; Schoenholz, S. S.; Aykol, M.; Cheon, G.; Cubuk, E. D. Scaling deep learning for materials discovery. *Nature* **2023**, *624*, 80—85.
- (4) Szymanski, N. J.; Rendy, B.; Fei, Y.; Kumar, R. E.; He, T.; Milsted, D.; McDermott, M. J.; Gallant, M.; Cubuk, E. D.; Merchant, A.; others An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* **2023**, *624*, 86—91.
- (5) Anwar, J.; Leitold, C.; Peters, B. Solid–solid phase equilibria in the NaCl–KCl system. *The Journal of Chemical Physics* **2020**, *152*.
- (6) Huang, J.; Ouyang, B.; Zhang, Y.; Yin, L.; Kwon, D.-H.; Cai, Z.; Lun, Z.; Zeng, G.; Balasubramanian, M.; Ceder, G. Inhibiting collective cation migration in Li-rich cathode materials as a strategy to mitigate voltage hysteresis. *Nature Materials* **2023**, *22*, 353–361.
- (7) Keen, D. A.; Goodwin, A. L. The crystallography of correlated disorder. *Nature* **2015**, *521*, 303–309.
- (8) Austin, I.; Goodman, C.; Pengelly, A. New semiconductors with the chalcopyrite structure. *Journal of The Electrochemical Society* **1956**, *103*, 609.
- (9) Anderson, M. T.; Greenwood, K. B.; Taylor, G. A.; Poeppelmeier, K. R. B-cation arrangements in double perovskites. *Progress in solid state chemistry* **1993**, *22*, 197–233.
- (10) Graf, T.; Felser, C.; Parkin, S. S. Simple rules for the understanding of Heusler compounds. *Progress in solid state chemistry* **2011**, *39*, 1–50.
- (11) Klayman, J.; Ha, Y.-W. Confirmation, disconfirmation, and information in hypothesis testing. *Psychological review* **1987**, *94*, 211.

- (12) Bertrand, R.; Russell, B. *The Philosophy of Logical Atomism*; Routledge classics; Routledge, 2009.
- (13) Schlesinger, C.; Fitterer, A.; Buchsbaum, C.; Habermehl, S.; Chierotti, M. R.; Nervi, C.; Schmidt, M. U. Ambiguous structure determination from powder data: four different structural models of 4,11-difluoroquinacridone with similar X-ray powder patterns, fit to the PDF, SSNMR and DFT-D. *IUCrJ* **2022**, *9*, 406–424.
- (14) Serra, M.; Salagre, P.; Cesteros, Y.; Medina, F.; Sueiras, J. E. Design of NiO–MgO materials with different properties. *Phys. Chem. Chem. Phys.* **2004**, *6*, 858–864.
- (15) Yagi, S.; Ichikawa, Y.; Yamada, I.; Doi, T.; Ichitsubo, T.; Matsubara, E. Synthesis of Binary Magnesium–Transition Metal Oxides via Inverse Coprecipitation. *Japanese Journal of Applied Physics* **2013**, *52*, 025501.
- (16) Tsirelson, V. G.; Avilov, A. S.; Abramov, Y. A.; Belokoneva, E. L.; Kitaneh, R.; Feil, D. X-ray and Electron Diffraction Study of MgO. *Acta Crystallographica Section B* **1998**, *54*, 8–17.
- (17) Fabrication of 1D mesoporous NiO nano-rods as high capacity and long-life anode material for lithium ion batteries. *Journal of Alloys and Compounds* **2021**, *850*, 156755.
- (18) Taguchi, H.; Omori, S.; Nagao, M.; Kido, H.; Shimada, M. Crystal Structure and Magnetic Properties of  $(\text{Ni}_{1-x}\text{Mg}_x)_6\text{MnO}_8$ . *Journal of Solid State Chemistry* **1995**, *118*, 112–116.
- (19) Guo, Y.; Wei, Y.; Li, H.; Zhai, T. Layer structured materials for advanced energy storage and conversion. *Small* **2017**, *13*, 1701649.
- (20) Delmas, C.; Fouassier, C.; Hagemuller, P. Structural classification and properties of the layered oxides. *Physica B+ c* **1980**, *99*, 81–85.

- (21) Seibel, E. M.; Roudebush, J.; Wu, H.; Huang, Q.; Ali, M. N.; Ji, H.; Cava, R. Structure and magnetic properties of the  $\alpha$ -NaFeO<sub>2</sub>-type honeycomb compound Na<sub>3</sub>Ni<sub>2</sub>BiO<sub>6</sub>. *Inorganic chemistry* **2013**, *52*, 13605–13611.
- (22) Kim, S. H.; Kim, S. J.; Oh, S. M. Preparation of layered MnO<sub>2</sub> via thermal decomposition of KMnO<sub>4</sub> and its electrochemical characterizations. *Chemistry of Materials* **1999**, *11*, 557–563.
- (23) Gardner, J. S.; Gingras, M. J.; Greedan, J. E. Magnetic pyrochlore oxides. *Reviews of Modern Physics* **2010**, *82*, 53.
- (24) Cascales, C.; Rasines, I.; Casado, P. G.; Vega, J. The new pyrochlores Pb<sub>2</sub>(M<sub>0.5</sub>Sb<sub>1.5</sub>)O<sub>6.5</sub> (M= Al, Sc, Cr, Fe, Ga, Rh). *Materials research bulletin* **1985**, *20*, 1359–1365.
- (25) Cascales, C.; Alonso, J.; Rasines, I. The new pyrochlores Pb<sub>2</sub>(MSb)O<sub>6.5</sub> (M= Ti, Zr, Sn, Hf). *Journal of materials science letters* **1986**, *5*, 675–677.
- (26) Roy, A.; Berrie, B. H. A new lead-based yellow in the seventeenth century. *Studies in Conservation* **1998**, *43*, 160–165.
- (27) Marchetti, A.; Saniz, R.; Krishnan, D.; Rabbachin, L.; Nuyts, G.; De Meyer, S.; Verbeeck, J.; Janssens, K.; Pelosi, C.; Lamoen, D.; others Unraveling the role of lattice substitutions on the stabilization of the intrinsically unstable Pb<sub>2</sub>Sb<sub>2</sub>O<sub>7</sub> pyrochlore: Explaining the lightfastness of lead pyroantimonate artists' pigments. *Chemistry of Materials* **2020**, *32*, 2863–2873.
- (28) Guillemet-Fritsch, S.; Baudour, J.; Chanel, C.; Bouree, F.; Rousset, A. X-ray and neutron diffraction studies on nickel zinc manganite Mn<sub>2.35-x</sub>Ni<sub>0.65</sub>Zn<sub>x</sub>O<sub>4</sub> powders. *Solid State Ionics* **2000**, *132*, 63–69.

- (29) Gama, L.; Paiva-Santos, C. O.; Vila, C.; Lisboa-Filho, P. N.; Longo, E. Characterization of nickel doped  $\text{Zn}_7\text{Sb}_2\text{O}_{12}$  spinel phase using Rietveld refinement. *Powder Diffraction* **2003**, *18*, 219–223.
- (30) Oleś, A. Neutron diffraction study of the  $\text{ZnFe}_{2-x}\text{Cr}_x\text{O}_4$  series. *physica status solidi (a)* **3**, 569–577.
- (31) Yamamoto, T.; Otsubo, Y.; Nagase, T.; Kosuge, T.; Azuma, M. Synthesis and Structure of Vacancy-Ordered Perovskite  $\text{Ba}_6\text{Ta}_2\text{Na}_2\text{X}_2\text{O}_{17}$  (X= P, V): Significance of Structural Model Selection on Discovered Compounds. *chemarxiv* **2023**,
- (32) Knapp, M. C.; Woodward, P. M. A-site cation ordering in  $\text{AA}'\text{BB}'\text{O}_6$  perovskites. *Journal of Solid State Chemistry* **2006**, *179*, 1076–1085.
- (33) Ji, K.; Alharbi, K. N.; Solana-Madruga, E.; Moyo, G. T.; Ritter, C.; Attfield, J. P. Double Double to Double Perovskite Transformations in Quaternary Manganese Oxides. *Angewandte Chemie International Edition* **60**, 22248–22252.
- (34) Wang, Y.; Li, W.; Hu, G.; Peng, Z.; Cao, Y.; Gao, H.; Du, K.; Goodenough, J. B. Electrochemical performance of large-grained  $\text{NaCrO}_2$  cathode materials for Na-ion batteries synthesized by decomposition of  $\text{Na}_2\text{Cr}_2\text{O}_7 \cdot 2\text{H}_2\text{O}$ . *Chemistry of Materials* **2019**, *31*, 5214–5223.
- (35) Nord, A. G. Neutron diffraction studies of  $\text{NiCoP}_4\text{O}_{12}$  and  $\text{NiZnP}_4\text{O}_{12}$ . *Materials Research Bulletin* **1983**, *18*, 765–773.
- (36) Wei, D.; Yang, X.; Liu, Y.; Seo, H. J. Structural and optical properties of  $\text{Mg}_{1-x}\text{Mn}_x\text{P}_2\text{O}_6$  ( $x = 0-1.0$ ) magnesium metaphosphate. *Journal of the American Ceramic Society* **2023**, *106*, 4246–4260.
- (37) Daidouh, A.; Durio, C.; Pico, C.; Veiga, M.; Chouaibi, N.; Ouassini, A. Structural and

- electrical study of the alluaudites  $(\text{Ag}_{1-x}\text{Na}_x)_2\text{FeMn}_2(\text{PO}_4)_3$  ( $x=0, 0.5$  and  $1$ ). *Solid State Sciences* **2002**, *4*, 541–548.
- (38) Barth, S.; Olazcuaga, R.; Gravereau, P.; Le Flem, G.; Hagemuller, P.  $\text{Mg}_{0.5}\text{Ti}_2(\text{PO}_4)_3$  — a new member of the NASICON family with low thermal expansion. *Materials Letters* **1993**, *16*, 96–101.
- (39) Hidouri, M.; Sendi, N.; Wattiaux, A.; Amara, M. B. Structural study by X-ray diffraction and Mössbauer spectroscopy of a new synthetic iron phosphate:  $\text{K}_4\text{MgFe}_3(\text{PO}_4)_5$ . *Journal of Physics and Chemistry of Solids* **2008**, *69*, 2555–2558.
- (40) Norberg, S. T.; Sobolev, A. N.; Streltsov, V. A. Cation movement and phase transitions in KTP isostructures; X-ray study of sodium-doped KTP at 10.5K. *Acta Crystallographica Section B* **2003**, *59*, 353–360.
- (41) Lenaz, D.; Princivale, F. Crystal chemistry of clinopyroxenes from Oligo-Miocene volcanics of Montresta (Sardinia, Italy). *Mineralogy and Petrology* **2007**, *90*, 157–166.
- (42) Griesemer, S. D.; Ward, L.; Wolverton, C. High-throughput crystal structure solution using prototypes. *Physical Review Materials* **2021**, *5*, 105003.
- (43) Li, C.; Zhong, J. Highly efficient broadband near-infrared luminescence with zero-thermal-quenching in garnet  $\text{Y}_3\text{In}_2\text{Ga}_3\text{O}_{12}:\text{Cr}^{3+}$  phosphors. *Chemistry of Materials* **2022**, *34*, 8418–8426.
- (44) Britvin, S. N.; Murashko, M. N.; Krzhizhanovskaya, M. G.; Vlasenko, N. S.; Vereshchagin, O. S.; Vapnik, Y.; Bocharov, V. N. Crocobelonite,  $\text{CaFe}_2^{3+}(\text{PO}_4)_2\text{O}$ , a new oxyphosphate mineral, the product of pyrolytic oxidation of natural phosphides. *American Mineralogist* **2023**, *108*, 1973–1983.
- (45) Peplow, M. Robot chemist sparks row with claim it created new materials. *Nature News* **2023**,

- (46) Toby, B. H. R factors in Rietveld analysis: How good is good enough? *Powder Diffraction* **2006**, *21*, 67 – 70.
- (47) Meneghini, C.; Ray, S.; Liscio, F.; Bardelli, F.; Mobilio, S.; Sarma, D. Nature of “disorder” in the ordered double perovskite Sr<sub>2</sub>FeMoO<sub>6</sub>. *Physical Review Letters* **2009**, *103*, 046403.
- (48) Mayo, R. A.; Marczenko, K. M.; Johnson, E. R. Quantitative matching of crystal structures to experimental powder diffractograms. *Chemical Science* **2023**, *14*, 4777–4785.
- (49) Lunt, A.; Fakhruddin, H.; Pizzuto, G.; Longley, L.; White, A.; Rankin, N.; Clowes, R.; Alston, B.; Gigli, L.; Day, G. M.; Chong, S. Y.; Cooper, A. Modular, Multi-Robot Integration of Laboratories: An Autonomous Workflow for Solid-State Chemistry. *Chem. Sci.* **2024**, –.
- (50) Salgado, J. E.; Lerman, S.; Du, Z.; Xu, C.; Abdolrahim, N. Automated classification of big X-ray diffraction data using deep learning models. *npj Computational Materials* **2023**, *9*, 214.
- (51) Jaros, M. Electronic properties of semiconductor alloy systems. *Reports on Progress in Physics* **1985**, *48*, 1091.
- (52) Bellaiche, L.; Vanderbilt, D. Virtual crystal approximation revisited: Application to dielectric and piezoelectric properties of perovskites. *Phys. Rev. B* **2000**, *61*, 7877–7882.
- (53) Chen, C.; Wang, E. G.; Gu, Y. M.; Bylander, D. M.; Kleinman, L. Unexpected band-gap collapse in quaternary alloys at the group-III-nitride/GaAs interface: GaAlAsN. *Phys. Rev. B* **1998**, *57*, 3753–3756.

- (54) Ramer, N. J.; Rappe, A. M. Virtual-crystal approximation that works: Locating a compositional phase boundary in  $\text{Pb}(\text{Zr}_{1-x}\text{Ti}_x)\text{O}_3$ . *Phys. Rev. B* **2000**, *62*, R743–R746.
- (55) Kikuchi, R. A Theory of Cooperative Phenomena. *Phys. Rev.* **1951**, *81*, 988–1003.
- (56) Sanchez, J.; Ducastelle, F.; Gratias, D. Generalized cluster description of multicomponent systems. *Physica A: Statistical Mechanics and its Applications* **1984**, *128*, 334–350.
- (57) Zunger, A.; Wei, S.-H.; Ferreira, L. G.; Bernard, J. E. Special quasirandom structures. *Phys. Rev. Lett.* **1990**, *65*, 353–356.
- (58) Grau-Crespo, R.; Hamad, S.; Catlow, C. R. A.; de Leeuw, N. H. Symmetry-adapted configurational modelling of fractional site occupancy in solids. *Journal of Physics: Condensed Matter* **2007**, *19*, 256201.
- (59) Yaghoobi, M.; Alaei, M. Machine learning for compositional disorder: A comparison between different descriptors and machine learning frameworks. *Computational Materials Science* **2022**, *207*, 111284.
- (60) Mazin, I. Inverse Occam’s razor. *Nature Physics* **2022**, *18*, 367–368.
- (61) Stokes, H. T. ; Campbell, B. J.; Hatch, D. M. FINDSYM. 2023; <https://stokes.byu.edu/iso/findsym.php>.

# Appendix

## Indexing of New Phases

Table S3: Table of new compounds synthesized by A-Lab compared against potential matches found in the ICSD database. The proposed symmetry is the symmetry of the proposed new phases as reported by A-lab, and the indexed symmetry is the symmetry found from FINDSYM<sup>61</sup> of the proposed phase using the provided structure. Discrepancies between the two symmetries are highlighted in pink. Some phases indexed using the DIFFRACT.EVA software are not tabulated in ICSD; these phases are denoted by (EVA). Note that not all of the mentioned ICSD phases have been refined against the provided PXRD pattern, and that there may be impurity phases that have yet to be identified.

Sample Candidates					
Proposed Phase	Proposed Symmetry	Indexed Symmetry	ICSD Phase	ICSD Code	ICSD Symmetry
Ba <sub>2</sub> ZrSnO <sub>6</sub>	Fm $\bar{3}m$ (225)	Pm $\bar{3}m$ (221)	BaSnO <sub>3</sub> Ba(Zr <sub>0.5</sub> Sn <sub>0.5</sub> O <sub>3</sub> ) SnO <sub>2</sub> ZrO <sub>2</sub>	188149 43137 9163 66781	Pm $\bar{3}m$ (221) Pm $\bar{3}m$ (221) P4 <sub>2</sub> /mmm (136) P4 <sub>2</sub> /nmc (137)
Ba <sub>6</sub> Na <sub>2</sub> Ta <sub>2</sub> V <sub>2</sub> O <sub>17</sub>	P6 <sub>3</sub> mmc (194)	P6 <sub>3</sub> mmc (194)	BaNaVO <sub>4</sub> NaVO <sub>3</sub>	130002 2103	P $\bar{3}m$ 1 (164) C12/c1 (15)
Ba <sub>6</sub> Na <sub>2</sub> V <sub>2</sub> Sb <sub>2</sub> O <sub>17</sub>	P6 <sub>3</sub> mmc (194)	P6 <sub>3</sub> mmc (194)	BaNaVO <sub>4</sub> NaVO <sub>3</sub>	130002 2103	P $\bar{3}m$ 1 (164) C12/c1 (15)
Ba <sub>9</sub> Ca <sub>3</sub> La <sub>4</sub> (Fe <sub>4</sub> O <sub>15</sub> ) <sub>2</sub>	P1 (1)	P1 (1)	Ba <sub>4.5</sub> Ca <sub>1.5</sub> La <sub>2</sub> Fe <sub>4</sub> O <sub>15</sub>	72336	P6 <sub>3</sub> mc (186)
CaCo(PO <sub>3</sub> ) <sub>4</sub>	C2/c (15)	C2/c (15)	Co <sub>2</sub> P <sub>2</sub> O <sub>7</sub> Ca <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub>	2830 923	B12 <sub>1</sub> /c1 (14) P12 <sub>1</sub> /a1 (14)
CaFe <sub>2</sub> P <sub>2</sub> O <sub>9</sub>	Pnma (62)	Pnma (62)	No matches	-	-
CaGd <sub>2</sub> Zr(GaO <sub>3</sub> ) <sub>4</sub>	P $\bar{1}$ (2)	P $\bar{1}$ (2)	Ca <sub>0.95</sub> Zr <sub>0.95</sub> — Gd <sub>2.05</sub> Ga <sub>4.05</sub> O <sub>12</sub>	202850	Ia $\bar{3}d$ (230)
CaMn(PO <sub>3</sub> ) <sub>4</sub>	C2/c (15)	C2/c (15)	Mn <sub>2</sub> P <sub>4</sub> O <sub>12</sub>	412558	C12/c1 (15)
CaNi(PO <sub>3</sub> ) <sub>4</sub>	C2/c (15)	C2/c (15)	CaNi <sub>3</sub> (P <sub>2</sub> O <sub>7</sub> ) <sub>2</sub> Ca <sub>2</sub> (P <sub>2</sub> O <sub>7</sub> )	74046 14313	P12 <sub>1</sub> /c1 (14) P4 <sub>1</sub> (76)
FeSb <sub>3</sub> Pb <sub>4</sub> O <sub>13</sub>	R3m (160)	R3m (160)	Pb <sub>2</sub> (Fe,Sb)O <sub>6.5</sub>	60805	Fd $\bar{3}m$ (227)

Sample Candidates (cont'd)					
Proposed Phase	Proposed Symmetry	Indexed Symmetry	ICSD Phase	ICSD Code	ICSD Symmetry
Hf <sub>2</sub> Sb <sub>2</sub> Pb <sub>4</sub> O <sub>13</sub>	Imm2 (44)	Imm2 (44)	Pb <sub>2</sub> (Hf,Sb)O <sub>6.5</sub> PbHfO <sub>3</sub>	62723 174110	Fd $\bar{3}m$ (227) Pbam (55)
InSb <sub>3</sub> (PO <sub>4</sub> ) <sub>6</sub>	Pc (7)	Pc (7)	Sb <sub>0.5</sub> In <sub>0.5</sub> (P <sub>2</sub> O <sub>7</sub> ) Sb(Sb <sub>0.5</sub> In <sub>0.5</sub> )(PO <sub>4</sub> ) <sub>3</sub>	166834 166835	P12 <sub>1</sub> /n1 (14) Pna2 <sub>1</sub> (33)
InSb <sub>3</sub> Pb <sub>4</sub> O <sub>13</sub>	R3m (160)	R3m (160)	Pb <sub>2</sub> (InSb)O <sub>6.5</sub> In <sub>2</sub> O <sub>3</sub>	41119 14388	Fd $\bar{3}m$ (227) Ia $\bar{3}$ (206)
K <sub>2</sub> TiCr(PO <sub>4</sub> ) <sub>3</sub>	P2 <sub>1</sub> 3 (198)	P2 <sub>1</sub> 3 (198)	K <sub>2</sub> Ti <sub>2</sub> (PO <sub>4</sub> ) <sub>3</sub> Cr <sub>2</sub> O <sub>3</sub>	202888 626479	P2 <sub>1</sub> 3 (198) R $\bar{3}c$ (167)
K <sub>4</sub> MgFe <sub>3</sub> (PO <sub>4</sub> ) <sub>5</sub>	Cc (9)	Cc (9)	K <sub>4</sub> MgFe <sub>3</sub> (PO <sub>4</sub> ) <sub>5</sub>	161484	P $\bar{4}2_1c$ (114)
K <sub>4</sub> TiSn <sub>3</sub> (PO <sub>5</sub> ) <sub>4</sub>	P2 <sub>1</sub> (4)	P2 <sub>1</sub> (4)	K((Ti <sub>0.25</sub> Sn <sub>0.75</sub> )O)(PO <sub>4</sub> )	250088	Pna2 <sub>1</sub> (33)
KBaGdWO <sub>6</sub>	F $\bar{4}3m$ (216)	F $\bar{4}3m$ (216)	Ba <sub>2</sub> GdWO <sub>6</sub> Gd <sub>2</sub> O <sub>3</sub>	138973 40473	Fm $\bar{3}m$ (225) Ia $\bar{3}$ (206)
KBaPrWO <sub>6</sub>	F $\bar{4}3m$ (216)	F $\bar{4}3m$ (216)	Ba <sub>11</sub> W <sub>4</sub> O <sub>23</sub> Pr <sub>6</sub> O <sub>11</sub> (EVA)	418207 N/A	Fd $\bar{3}m$ (227) P2 <sub>1</sub> /c (14)
KMn <sub>3</sub> O <sub>6</sub>	C2/c (15)	C2/m (12)	K <sub>0.48</sub> Mn <sub>1.94</sub> O <sub>5.18</sub>	240249	P6 <sub>3</sub> /mmc (194)
KNa <sub>2</sub> Ga <sub>3</sub> (SiO <sub>4</sub> ) <sub>3</sub>	P2 <sub>1</sub> /c (14)	P2 <sub>1</sub> /c (14)	NaGaSiO <sub>4</sub>	46861	P12 <sub>1</sub> /n1 (14)
KNaP <sub>6</sub> (PbO <sub>3</sub> ) <sub>8</sub>	P3 (143)	P6 <sub>3</sub> /m (176)	(Na,K)Pb <sub>4</sub> (PO <sub>4</sub> ) <sub>3</sub> Pb <sub>8</sub> O <sub>5</sub> (PO <sub>4</sub> ) <sub>2</sub>	182501 98702	P6 <sub>3</sub> /m (176) C12/m1 (12)
KNaTi <sub>2</sub> (PO <sub>5</sub> ) <sub>2</sub>	Pna2 <sub>1</sub> (33)	Pna2 <sub>1</sub> (33)	Na <sub>0.5</sub> (Na <sub>0.492</sub> K <sub>0.008</sub> )– (TiO)(PO <sub>4</sub> )	59284	Pna2 <sub>1</sub> (33)
KPr <sub>9</sub> (Si <sub>3</sub> O <sub>13</sub> ) <sub>2</sub>	P3 (143)	P3 (143)	Pr <sub>9.33</sub> Si <sub>6</sub> O <sub>32</sub> (EVA)	N/A	-
Mg <sub>3</sub> MnNi <sub>3</sub> O <sub>8</sub>	R $\bar{3}m$ (166)	R $\bar{3}m$ (166)	NiO (Ni,Mn)(Ni,Mn) <sub>2</sub> O <sub>4</sub>	9866 84517	Fm $\bar{3}m$ (225) Fd $\bar{3}m$ (227)
Mg <sub>3</sub> NiO <sub>4</sub>	Pm $\bar{3}m$ (221)	Fm $\bar{3}m$ (225)	MgNiO <sub>2</sub>	290603	Fm $\bar{3}m$ (225)
MgCuP <sub>2</sub> O <sub>7</sub>	P $\bar{1}$ (2)	P $\bar{1}$ (2)	Mg <sub>2</sub> P <sub>2</sub> O <sub>7</sub>	20295	C12/m1 (12)
MgNi(PO <sub>3</sub> ) <sub>4</sub>	C2/c (15)	C2/c (15)	Mg <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> Ni <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> (Ni,Mn) <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub>	31005 153159 158525	P12 <sub>1</sub> /n1 (14) P12 <sub>1</sub> /c1 (14) P12 <sub>1</sub> /a1 (14)
MgTi <sub>2</sub> NiO <sub>6</sub>	R3 (146)	R3 (146)	(Ni <sub>0.5</sub> Mg <sub>0.5</sub> )TiO <sub>3</sub> TiO	171583 38755	R $\bar{3}$ (148) Fm $\bar{3}m$ (225)
MgTi <sub>4</sub> (PO <sub>4</sub> ) <sub>6</sub>	R3 (146)	R3 (146)	Mg <sub>0.5</sub> Ti <sub>2</sub> (PO <sub>4</sub> ) <sub>3</sub>	74287	R $\bar{3}c$ (167)
MgV <sub>4</sub> Cu <sub>3</sub> O <sub>14</sub>	P1 (1)	P1 (1)	(Cu <sub>1.5</sub> Mg <sub>0.5</sub> )V <sub>2</sub> O <sub>7</sub>	69731	C12/c1 (15)

Sample Candidates (cont'd)					
Proposed Phase	Proposed Symmetry	Indexed Symmetry	ICSD Phase	ICSD Code	ICSD Symmetry
Mn <sub>2</sub> VPO <sub>7</sub>	Cm (8)	Cm (8)	Mg <sub>2</sub> P <sub>2</sub> O <sub>7</sub>	47136	C12/m1 (12)
Mn <sub>4</sub> Zn <sub>3</sub> (NiO <sub>6</sub> ) <sub>2</sub>	C2/c (15)	C2/c (15)	(Zn <sub>0.759</sub> Mn <sub>0.241</sub> )– (Mn <sub>1.35</sub> Ni <sub>0.65</sub> )O <sub>4</sub> NiO	92223 9866	Fd $\bar{3}$ m (227) Fm $\bar{3}$ m (225)
Mn <sub>7</sub> (P <sub>2</sub> O <sub>7</sub> ) <sub>4</sub>	C222 <sub>1</sub> (20)	C222 <sub>1</sub> (20)	Mn <sub>2</sub> (PO <sub>3</sub> ) <sub>4</sub>	145534	C12/c1 (15)
MnAgO <sub>2</sub>	C2/m (12)	C2/m (12)	MnAgO <sub>2</sub>	139006	C12/m1 (12)
Na <sub>3</sub> Ca <sub>18</sub> Fe(PO <sub>4</sub> ) <sub>14</sub>	P1 (1)	P1 (1)	Na <sub>3</sub> Ca <sub>18</sub> Fe(PO <sub>4</sub> ) <sub>14</sub>	85103	R3c (161)
Na <sub>7</sub> Mg <sub>7</sub> Fe <sub>5</sub> (PO <sub>4</sub> ) <sub>12</sub>	P1 (1)	P1 (1)	Na <sub>2</sub> Mg <sub>2</sub> Fe(PO <sub>4</sub> ) <sub>3</sub>	138263	C12/c1 (15)
NaCaMgFe(SiO <sub>3</sub> ) <sub>4</sub>	C2 (5)	C2 (5)	(Ca <sub>0.774</sub> Na <sub>0.226</sub> )– (Mg <sub>0.901</sub> Fe <sub>0.099</sub> )– Fe <sub>0.011</sub> (Si <sub>2</sub> O <sub>6</sub> )	75294	C12/c1 (15)
NaMnFe(PO <sub>4</sub> ) <sub>2</sub>	P $\bar{1}$ (2)	P $\bar{1}$ (2)	Na <sub>2</sub> Mg <sub>2</sub> Fe(PO <sub>4</sub> ) <sub>3</sub>	138263	C12/c1 (15)
Sn <sub>2</sub> Sb <sub>2</sub> Pb <sub>4</sub> O <sub>13</sub>	Imm2 (44)	Imm2 (44)	Pb <sub>2</sub> SnSbO <sub>6.5</sub>	62722	Fd $\bar{3}$ m (227)
Y <sub>3</sub> In <sub>2</sub> Ga <sub>3</sub> O <sub>12</sub>	Ia $\bar{3}$ d (230)	Ia $\bar{3}$ d (230)	Y <sub>3</sub> In <sub>2</sub> Ga <sub>3</sub> O <sub>12</sub>	54664	Ia $\bar{3}$ d (230)
Zn <sub>2</sub> Cr <sub>3</sub> FeO <sub>8</sub>	R $\bar{3}$ m (166)	R $\bar{3}$ m (166)	(Zn <sub>0.54</sub> Fe <sub>0.46</sub> )Fe <sub>2</sub> O <sub>4</sub> Zn(FeCrO <sub>4</sub> )	81207 167362	Fd $\bar{3}$ m (227) Fd $\bar{3}$ m (227)
Zn <sub>3</sub> Ni <sub>4</sub> (SbO <sub>6</sub> ) <sub>2</sub>	C2/c (15)	C2/c (15)	Zn(Zn <sub>1.333</sub> Sb <sub>0.667</sub> )O <sub>4</sub> NiO NiSb <sub>2</sub> O <sub>6</sub>	173996 9866 426852	Fd $\bar{3}$ m (227) Fm $\bar{3}$ m (225) P4 <sub>2</sub> /mnm (136)
Zr <sub>2</sub> Sb <sub>2</sub> Pb <sub>4</sub> O <sub>13</sub>	Imm2 (44)	Imm2 (44)	Pb <sub>2</sub> (ZrSb)O <sub>6.5</sub>	62721	Fd $\bar{3}$ m (227)

## Additional Refinements

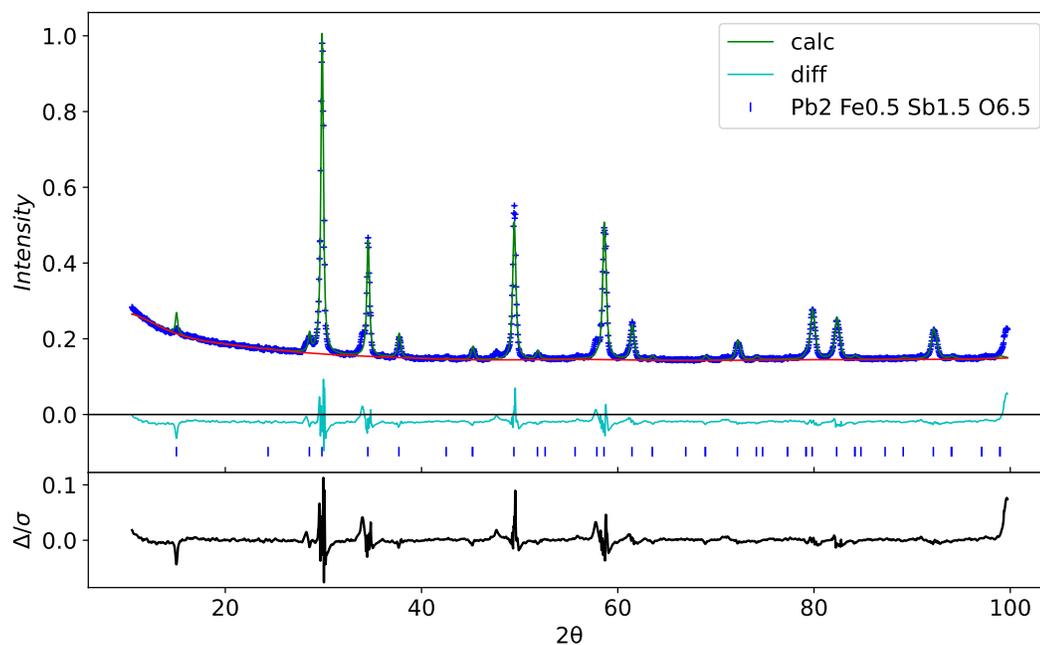


Figure S11: Rietveld refinement of the pattern representing  $\text{FeSb}_3\text{Pb}_4\text{O}_{14}$  against  $\text{Pb}_2\text{Fe}_{0.5}\text{Sb}_{1.5}\text{O}_{6.5}$  (coll-60805) as reported on ICSD.

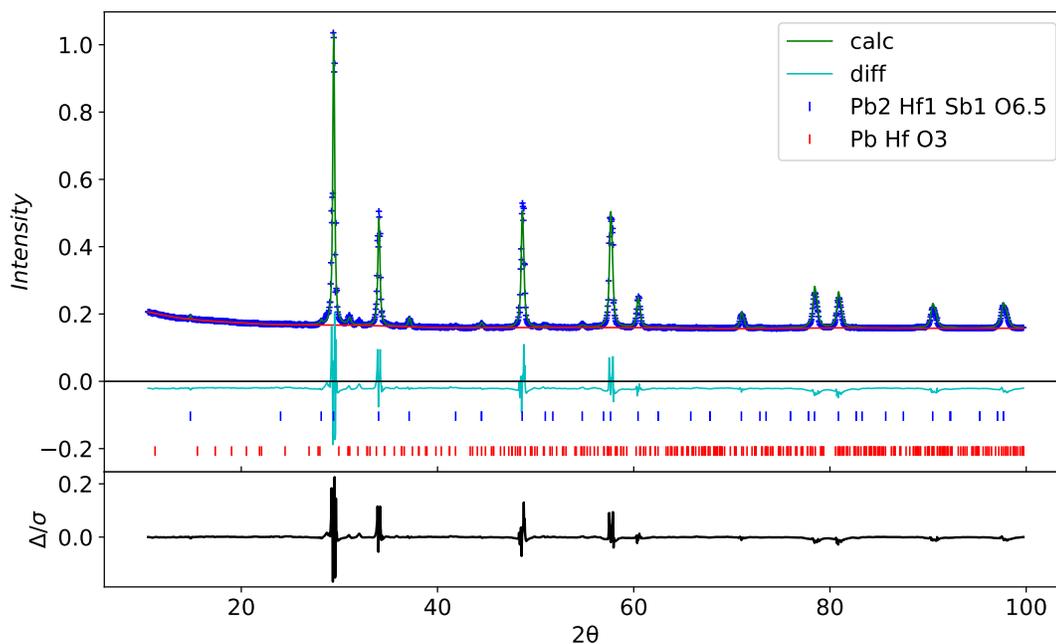


Figure S12: Rietveld refinement of the pattern representing  $\text{Hf}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$  against  $\text{Pb}_2\text{HfSbO}_{6.5}$  (coll-60805) and  $\text{PbHfO}_3$  (coll-174110) as reported on ICSD.

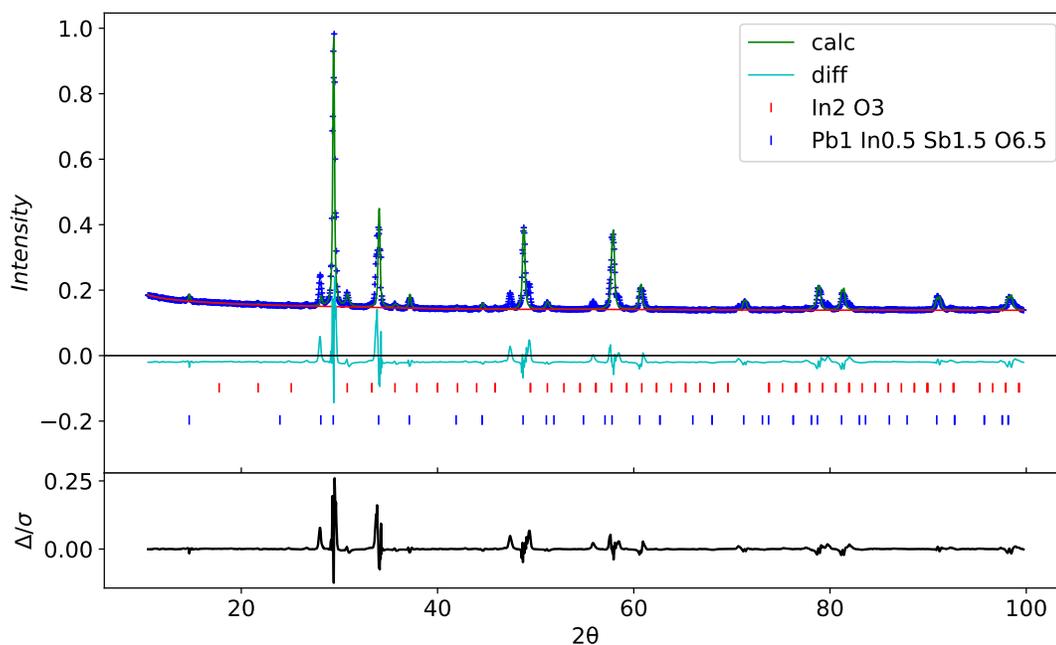


Figure S13: Rietveld refinement of the pattern representing  $\text{InSb}_3\text{Pb}_4\text{O}_{13}$  against  $\text{Pb}_1\text{In}_{0.5}\text{Sb}_{1.5}\text{O}_{6.5}$  (coll-14388) and  $\text{In}_2\text{O}_3$  (coll-41119) as reported on ICSD.

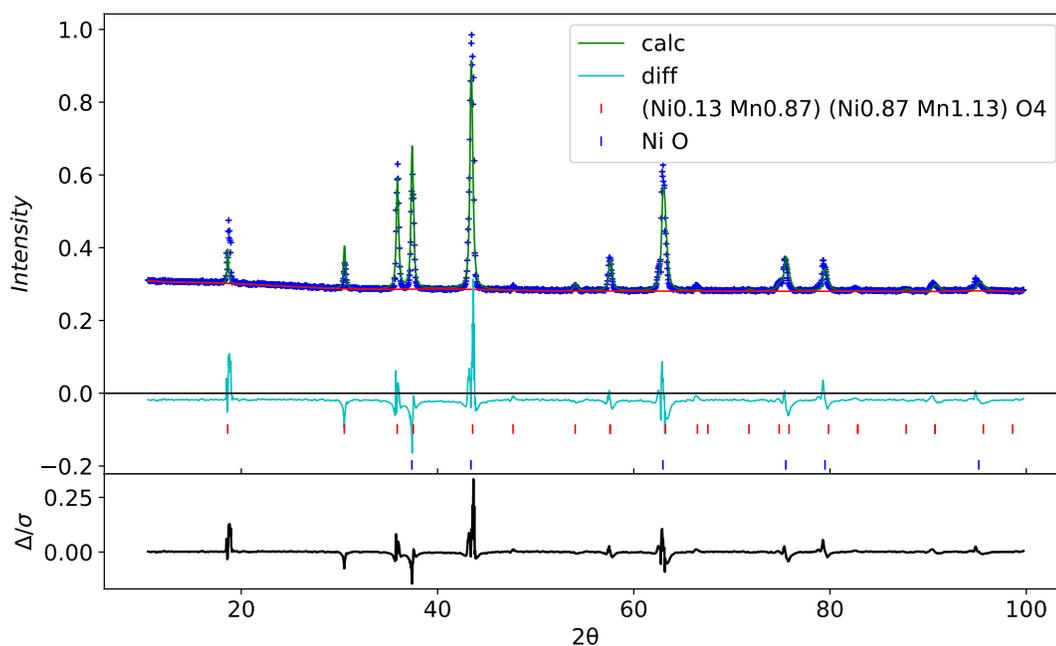


Figure S14: Rietveld refinement of the pattern representing  $\text{Mg}_3\text{MnNi}_3\text{O}_8$  against  $(\text{Ni}_{0.13}\text{Mn}_{0.87})(\text{Ni}_{0.87}\text{Mn}_{1.13})\text{O}_4$  (coll-84517) and  $\text{NiO}$  (coll-9866) as reported on ICSD.

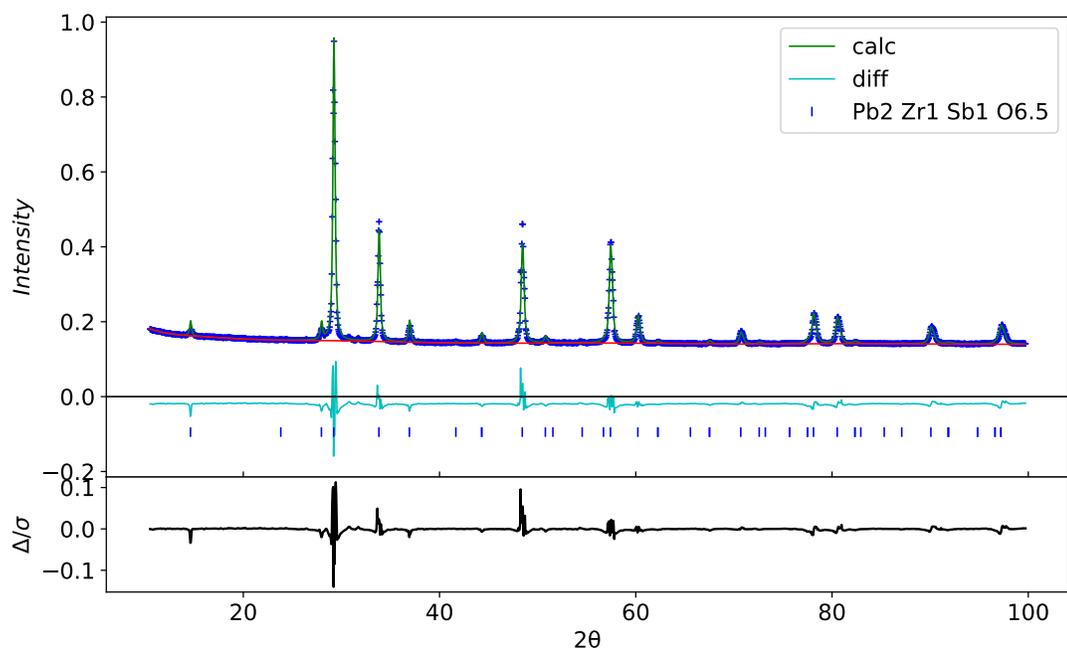


Figure S15: Rietveld refinement of the pattern representing  $\text{Zr}_2\text{Sb}_2\text{Pb}_4\text{O}_{13}$  against  $\text{Pb}_2\text{Zr}_1\text{Sb}_1\text{O}_{6.5}$  (coll-62721) as reported on ICSD.