Perspective

Challenges in High-Throughput Inorganic Materials Prediction and Autonomous Synthesis

Josh Leeman,[1] Yuhan Liu,[2] Joseph Stiles,[1,3] Scott B. Lee,[1] Prajna Bhatt,[2] Leslie M. Schoop,[1,3,*] and Robert G. Palgrave[2,†]

[1] *Department of Chemistry, Princeton University, Princeton, New Jersey 08540, USA*
[2] *Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, United Kingdom*
[3] *Princeton Materials Institute, Princeton University, Princeton, New Jersey 08540, USA*

(Received 30 January 2024; published 7 March 2024)

Materials discovery lays the foundation for many technological advancements. The prediction and discovery of new materials are not simple tasks. Here, we 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.* [Nature 624, 86 (2023)], 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, as follows. (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 for autonomous materials discovery but also for 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.

DOI: 10.1103/PRXEnergy.3.011002

I. 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) [1], 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, e.g., 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, the testing of new compositions, the determination of 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 [2] has been one approach to expanding the space of inorganic materials. It catalogs the known ICSD

*lschoop@princeton.edu
†r.palgrave@ucl.ac.uk

compounds with additional computational information and also suggests computationally predicted new materials.

Very recently, Google DeepMind have reported the prediction of up to 2.2 million new stable inorganic crystals, tabulated in their GNoME database [3]. However, some of these predictions may warrant experimental verification and synthesizing so many material candidates by hand would be extremely laborious. To accelerate this, a group at Berkeley has established an automated lab, called A-lab, that utilizes robotics and artificial intelligence. 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 has been successful, and if not, adjusts the synthetic conditions. The group behind A-lab has recently reported that within 17 days, A-lab has been able synthesize 41 new materials out of 58 predicted targets, an impressive success rate of 71% [4]. Using human intervention, the success rate has been 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 Perspective, 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 artificial intelligence (AI) can predict reasonable synthetic routes based on literature precedent, and that a full circle of materials synthesis and characterization can be carried out without human intervention. Unfortunately, we have found that the central claim of the A-lab paper, namely, that a large number of previously unknown materials have been 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 GNoME's screenshot of the ICSD and the time of the A-lab publication, meaning that they would not be in the original training set.

We find that the vast majority of the synthetic products have been wrongly characterized. These misinterpreted characterizations broadly fall under two categories: either the authors have failed to recognize that the automated refinement process has changed the symmetry of the target compound or the PXRD pattern agrees better with known phases—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 of 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." We highlight that any work that claims to have discovered new materials must define what makes those new, especially in relation to already known ones. As we hope to reach a multidisciplinary audience, we will then go into some, but not all, of the 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 that we see as motifs arising in each class.

II. 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. Nonmolecular 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, nonmolecular materials do not have a chemical composition restricted by molecular formulas. Their composition is not quantized but, instead, can be incrementally changed. One primary example of this is solid solutions, e.g., a solution of KCl and NaCl, which could be written $Na_{1-x}K_xCl$, where x can take any value between 0 and 1 [5]. 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 materials 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 crystal 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 has recently been demonstrated in $Li_{1.2}Cr_{0.4}Mn_{0.4}O_2$ [6]. Another example is the aforementioned $Na_{1-x}K_xCl$ 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 $Na_{1-x}K_xCl$. 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. First, it is commensurate with the experimental diffraction patterns—the PXRD patterns of $Na_{1-x}K_xCl$ solid solutions closely resemble 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. Second, 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 in, e.g., Rietveld refinement, while simple heuristics such as 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 transitional symmetry [7]. 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 chalchopyrite structure when the cations are ordered [8]. Chalchopyrite ($CuFeSe_2$) 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 [9]. In the case of alloys, Heusler alloys are a common example of ordered intermetallic compounds [10].

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 the 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 can 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, it would have diffraction patterns intermediate between those of 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 the context or field. We would wish to highlight that 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 the Materials Project and the larger Google GNoMe, fall in the category of structurally new materials. We have certainly not looked through all predicted new materials in the GNoMe data set—not even a large fraction. However, in the synthetic targets in the A-lab paper [4], 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 aristotype 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. We would also state that we have noticed similar issues in the GNoMe database when only looking at a small fraction of the predicted materials.

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 presence in the Materials Project (MP) and its absence from the ICSD. This criterion is open to criticism; e.g., MP contains no compositionally disordered

compounds and many known compounds are not in the ICSD, especially disordered ones) but, nonetheless, we will mostly use this criterion to assess the novelty of the A-lab synthesis products.

III. 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 [11]. 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 of forming 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, such as XY_2Z_4. Likewise, the unreacted starting materials should also be eliminated from enquiry and unintended contaminants ruled out. 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 has also been expressed in similar terms by Russell: "*Whenever possible, substitute constructions out of known entities for inferences to unknown entities*" [12].

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 the PXRD pattern and the 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 that the modeled compound is the correct one. Schlesinger *et al.* [13] point out that the "$\cdots$ *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. The 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, e.g., 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 material properties. The standard set by the

TABLE I. The distribution of errors in the 36 claimed "successful" syntheses. The "X" symbol denotes that 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 has been used for refinement compared with that in the paper and in the 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_2ZrSnO_6	X	X	X		$Mg_3MnNi_3O_8$	X		X	
$Ba_6Na_2Ta_2V_2O_{17}$	X		X		Mg_3NiO_4		X	X	
$Ba_6Na_2V_2Sb_2O_{17}$	X				$MgCuP_2O_7$		X	X	
$CaCo(PO_3)_4$			X		$MgNi(PO_3)_4$	X		X	
$CaFe_2P_2O_9$					$MgTi_2NiO_6$			X	
$CaMn(PO_3)_4$			X		$MgTi_4(PO_4)_6$				X
$CaNi(PO_3)_4$			X		$MgV_4Cu_3O_{14}$	X	X	X	
$FeSb_3Pb_4O_{13}$			X		Mn_2VPO_7	X		X	
$Hf_2Sb_2Pb_4O_{13}$			X		$Mn_4Zn_3(NiO_6)_2$			X	
$InSb_3Pb_4O_{13}$			X		$MnAgO_2$	X			X
$K_2TiCr(PO_4)_3$			X		$Na_3Ca_{18}Fe(PO_4)_{14}$	X			
$K_4MgFe_3(PO_4)_5$	X				$Na_7Mg_7Fe_5(PO_4)_{12}$	X			
$K_4TiSn_3(PO_5)_4$	X				$NaCaMgFe(SiO_3)_4$		X	X	
$KBaPrWO_6$	X				$NaMnFe(PO_4)_2$	X			
KMn_3O_6	X	X	X		$Sn_2Sb_2Pb_4O_{13}$			X	
$KNaP_6(PbO_3)_8$	X	X	X		$Y_3In_2Ga_3O_{12}$	X			X
$KNaTi_2(PO_5)_2$			X		$Zn_2Cr_3FeO_8$			X	
$KPr_9(Si_3O_{13})_2$	X	X			$Zr_2Sb_2Pb_4O_{13}$			X	

A-lab paper of $> 50\%$ purity being "success" is therefore anomalous in the usual practice of solid-state chemistry.

IV. ANALYSIS OF THE A-LAB DATA SET

Below, we go through the materials that have been claimed to be successfully synthesized in the A-lab paper [4]. We summarize our findings here, before going through many examples in detail. The classification of samples by 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 are "successes" (including two that were successful offline, meaning with human intervention in the synthesis).

Within the 36 samples classified as successes, we have found that the analysis presented for 35 of them has

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 or 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) *Changing the prediction to match the observed XRD pattern.* In several cases, the crystallographic information file (CIF) supplied in the Supplemental Information is not the same as that claimed in the main paper. The A-lab paper states "For targets for which we suspect the poor fit resulted from configurational disorder, we refined the XRD patterns using cation-disordered versions of the structure of the target taken from the Materials Project." In effect, the prediction has been changed to a disordered version of the compound if the ordered version did not fit. However, in the large majority of cases, as we show below, the *disordered version actually exists already in the ICSD*, meaning that the material is in fact already known. An example is Mg_3NiO_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 that are ordered versions of known disordered compounds. For example, as we will show in detail below, the existence of $MgTi_2NiO_6$ 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 compound. 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 Supplemental Information of the A-lab paper using GetData Graph Digitalizer. Because the software had trouble identifying the green dots that represented the observed XRD data, the experimental data were obtained by combining the calculated fit with the residual of the fit. To align the x-axis values for combining, the acquired XY data were then interpolated with the data in ORIGIN. This process is certainly not ideal and yields data of lower quality than the original. Nevertheless, we have found that it is possible to carry out Rietveld refinement on these data sets. This has been carried out in GSAS-II, the software used by A-lab [14]. We do not have the experimental parameters for the original data collection and so peak profiles have been determined empirically. We do not claim that our fits are definitive or cannot be improved upon but we highlight in each case the features that make us believe that the fits we propose are superior to those provided in the original paper. We note that the scale of the Δ/σ subplots of our refinements is arbitrary.

A. Rock-salt structured materials

As mentioned in Sec. I, 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.

Mg_3NiO_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, $MgNiO_2$, has been reported to exist in a rock-salt structure (space group $Fm\bar{3}m$, ICSD-290603), where no cation order was observed [see Fig. 1(b)] The sample synthesized by A-lab with the composition Mg_3NiO_4 has been claimed to be a successful synthesis after human intervention 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 pattern can be relatively well indexed with space group $Fm\bar{3}m$ [as shown in the Supplemental Information in the paper and also in Fig. 1(c)] but lacks additional peaks (e.g., the 100 peaks around 21.1°) 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 $k + l = 2n$ agree very well with the known compound $MgNiO_2$, as shown in Fig. 1; however, the peaks appear at slightly different diffraction angles, which may suggest a doped material or a solid

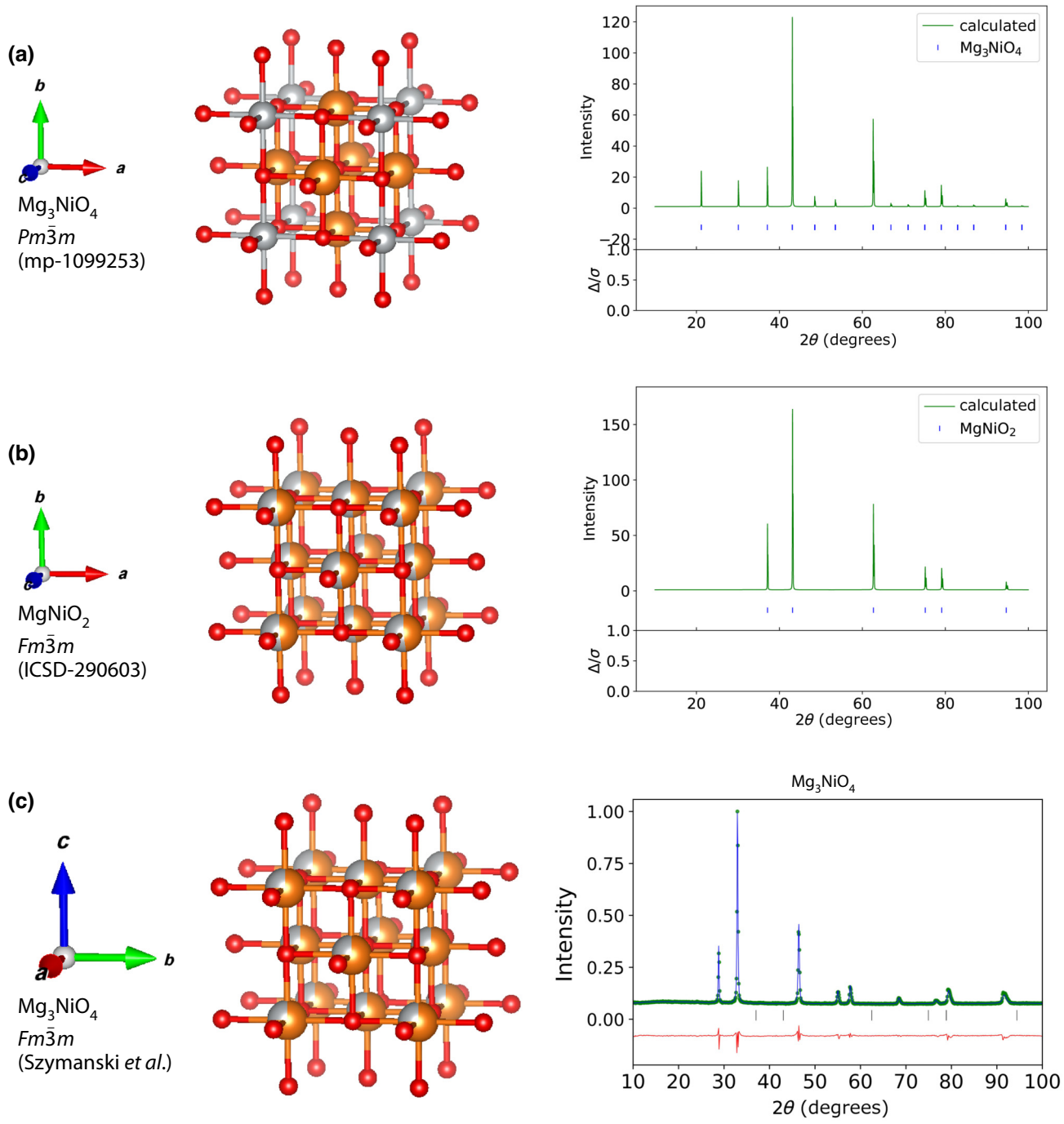

FIG. 1. (a) The structure of Mg_3NiO_4 as predicted by the Materials Project (mp-10999253, $Pm\bar{3}m$; left) and the simulated PXRD pattern of the same structure (right). (b) The structure of $Mg_2Ni_2O_4$ as listed in the ICSD (ICSD-290603, $Fm\bar{3}m$; left) and the simulated PXRD pattern of the same structure (right). (c) The structure of Mg_3NiO_4 as provided by Szymanski *et al.* [4] ($Fm\bar{3}m$; left) and the measured powder pattern given in the same paper (right). Mg is shown in orange, Ni in gray, and O in red.

solution. MgO-NiO solid solutions are well studied as catalytic materials and solid solutions can be formed across the composition range [15]. $Mg_2Ni_2O_4$ is reported with lattice parameter 4.1889(1) Å [16], whereas the CIF of Mg_3NiO_4 provided in the A-lab paper indicates a slightly larger lattice parameter of 4.20311 Å. By interpolating between the lattice parameters of rock salt MgO (4.214 Å) [17] and (metastable) rock salt NiO (4.1718 Å) [18], Vegard's law places the composition of a solid solution with the lattice parameter 4.20311 Å at exactly Mg_3NiO_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 synthesized 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 the 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 Mg_3NiO_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 Mg_3NiO_4 suffers from errors 2 (different structure used in refinement than was predicted) and 3 (no evidence for cation order in the predicted structure).

$Mg_3MnNi_3O_8$ 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 (ICSD-80306), reported by Taguchi *et al.* in 1995 [19]. The reported compound is cubic, a variant of the rock-salt structure, sometimes called murdochite, with octahedrally coordinated metals; the Mn ions form an 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 that 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°. In the murdochite phase ($Fm\bar{3}m$), the (111) peak at 18.4° 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°. This suggests that the predicted ordering is not present. The generally poor agreement in intensities may point to multiple phases being 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 unmodeled peak in the experimental pattern at just over 30°. 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 known phases: rock salt, spinel, and murdochite, with the Mg, Mn, and Ni possibly distributed across all these phases. The evidence from the peak intensities is clearly against the proposed Ni-Mg ordering. See Fig. 14 in the appendix for our alternative refinement using the phases described above. The analysis of $Mg_3MnNi_3O_8$ sufferers from errors 1 (very poor fit) and 3 (no evidence of cation order).

B. Layered materials

Layered materials are of significant interest in materials science, as they provide the foundation of many applications, including most battery-electrode materials [20]. Among the 43 materials that A-lab has synthesized, there is one layered compound, KMn_3O_6. This compound has been predicted by the Materials Project (mp-1016190) to crystallize in space group $C2/c$. The structure can be viewed as related to α-$NaFeO_2$, which consists of layers of edge-sharing 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 [21]. Cation order in the transition-metal layer can lower the symmetry to the monoclinic space group $C2/m$ [22]. The proposed structure of KMn_3O_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 Supplemental Information of the A-lab paper [4] 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 $K_{0.33}MnO_2$. Intuitively, one would expect that $K_{0.33}MnO_2$ would adopt one of the α-$NaFeO_2$-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 Figs. 2(a) and 2(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 [4]. For example, $K_{0.3}MnO_2$ has been reported in the hexagonal space group $P6_3/mmc$, in a structure that belongs to one of the stacking variants of the α-$NaFeO_2$-structure [23]. A compound of almost exactly the same stoichiometry is in the ICSD (ICSD-156081). This structure is shown in Fig. 2(c); it has neither ordered K nor buckled layers. $K_{0.3}MnO_2$ is known to result from the thermal decomposition of $KMnO_4$ above 800°C [23]. As A-lab's reaction conditions have included a 1000°C heating step, $K_{0.3}MnO_2$ is a likely product. The PXRD fit for KMn_3O_6 provided in the A-lab paper [4] is of very poor quality and misses some major reflections, as shown in Fig. 2(d). Comparing it to the simulated PXRD pattern of $K_{0.3}MnO_2$ [Fig. 2(e)] reveals that the main measured reflections are well reproduced by those simulated to appear $K_{0.3}MnO_2$. Still, there

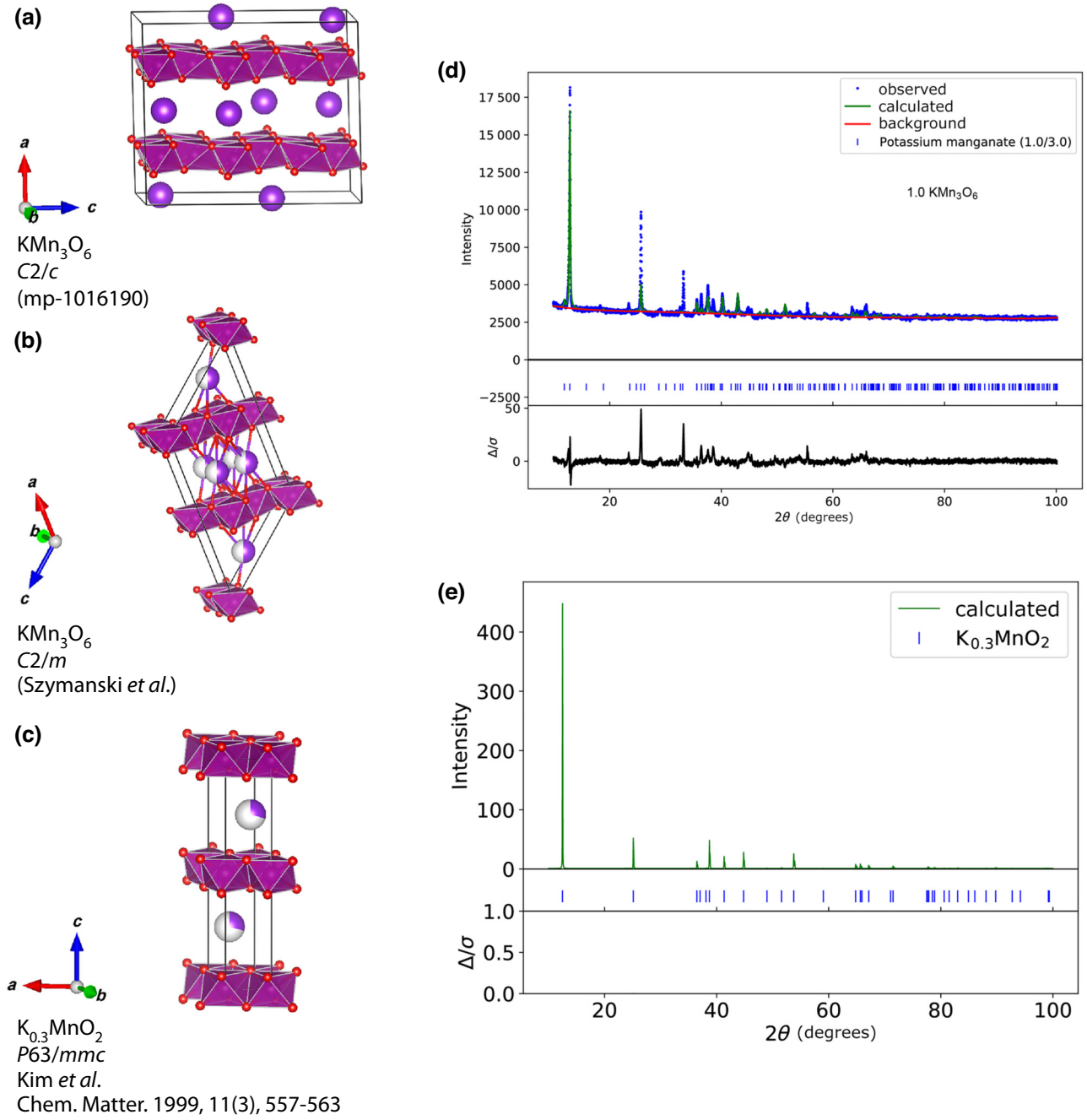

FIG. 2. (a) The structure of KMn_3O_6 as predicted by the Materials Project (mp-1016190, $C2/c$). (b) The structure of KMn_3O_6 as provided by Szymanski *et al.* [4] ($C2/m$). (c) The structure of $K_{0.3}MnO_2$ as reported in Kim *et al.* [23] ($P63/mmc$). K is shown as large purple spheres, Mn is shown in pink octahedra, and O is shown in red. (d) The PXRD pattern as provided by Szymanski *et al.* (e) The simulated PXRD pattern for $K_{0.3}MnO_2$.

is an intensity mismatch, which could be caused by preferred orientation, 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 errors 1 (poor fit), 2 (CIF file not the same as originally predicted), and 3 (no evidence of cation order).

C. Pb-Sb pyrochlores

Pyrochlores are structures that have formula $A_2B_2O_{7-\delta}$, with A and B denoting two possible cation sites and δ the possible oxygen defect [24]. 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 nonstoichiometry: A-site vacancies or Pb(IV) defects on the B site are compensated by oxygen vacancies, i.e., $\delta > 0$. Furthermore, the B site in lead pyrochlores can be occupied by two different cation species; e.g., if half of the B(V) site is replaced by a M(IV) cation, the resulting formula is $Pb_4M_2Sb_2O_{13}$—these compounds are known for M(IV) = 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 $Pb_4M_1Sb_3O_{13}$; example M(III) cations that have been incorporated in this way are In, Al, Sc, Cr, Fe, Ga, and Rh [25,26].

The Materials Project predicts the stability of various compounds with the formula $Pb_4M_xSb_{4-x}O_{13}$, where $x = 1, 2$ and M is a cation. A-lab has reported the successful synthesis and refinement of five $Pb_4M_xSb_{4-x}O_{13}$ compounds, namely, M = Fe, In, $x = 1$ ($FeSb_3Pb_4O_{13}$ and $InSb_3Pb_4O_{13}$), which are predicted to crystallize in space group $R3m$, as well as M = Hf, Sn, Zr, $x = 2$ ($Hf_2Sb_2Pb_4O_{13}$, $Sn_2Sb_2Pb_4O_{13}$, and $Zr_2Sb_2Pb_4O_{13}$), which are predicted to crystallize in space group Imm2. Note that 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 places 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 (ICSD-62722, 62721, 60805, 41119). In fact, all of the B sites in question have been 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 $Fd\bar{3}m$ [25,26].

Since the predicted structures for these five compounds are just B-site ordered versions of the known disordered structures, we again emphasize that evidence of the ordering must be found in the characterization 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 M = In, Hf, and Sn, meaning that any ordering of these elements in the synthesized materials would be very difficult to detect via XRD.

The ordered B-site cations in the predicted structures lead to a lower-symmetry unit cell compared with the cubic pyrochlore of the parent $Pb_2Sb_2O_{6.5}$. However, if the atom type is ignored, the atom positions are almost the same as in the cubic pyrochlore. We have calculated the pseudo-cubic lattice parameter, a_p, of the ordered CIFs produced by A-lab, as shown in Table II. Overall, these show close agreement with the known phases. The largest difference is the $Sn_2Sb_2Pb_4O_{13}$ compound, which we discuss in detail below. Other compounds have closely matching parameters and with no evidence of order we conclude that they are very likely the known disordered pyrochlore compounds discovered in the 1980s.

TABLE II. A comparison of the pseudo-cubic lattice parameters predicted in the 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 $Pb_2Sb_2O_7$	...	10.44
$Sn_2Sb_2Pb_4O_{13}$	10.6168	10.5645
$Zr_2Sb_2Pb_4O_{13}$	10.6594	10.6349
$Hf_2Sb_2Pb_4O_{13}$	10.6415	10.6126
$FeSb_3Pb_4O_{13}$	10.4931	10.4803
$InSb_3Pb_4O_{13}$	10.5845	10.5892

Figure 3 shows that the calculated diffraction pattern of the known phase $Pb_2SnSbO_{6.5}$ (ICSD-62722) and the predicted $Sn_2Sb_2Pb_4O_{13}$ 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 $Sn_2Sb_2Pb_4O_{13}$. And, most importantly, it needs to be verified that this material is different from $Pb_2SnSbO_{6.5}$ [27]. We find that the Rietveld refinement of the PXRD data from $Sn_2Sb_2Pb_4O_{13}$, shown in Fig. 4, can be carried out successfully using the known compounds $Pb_2SnSbO_{6.5}$ and SnO_2 (ICSD-9163), both reported in the 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 Fig. 4 and in Figs. 11, 12, 13, and 15 in the Appendix, indicate 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 3 (missed disorder or no evidence of cation order).

Notably, all these materials are related to the famous "Naples Yellow" pigment, which derives from $Pb_2Sb_2O_7$ [28]. 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 civilizations, in the Middle Ages, at various points in the Renaissance, and most recently by A-lab. Interestingly, it has been debated that $Pb_2Sb_2O_7$ itself is not stable and that doping (most commonly with Sn but also with other elements) is necessary to stabilize the pigment. For an excellent overview, we recommend the work of Marchetti *et al.* [28] and the references therein.

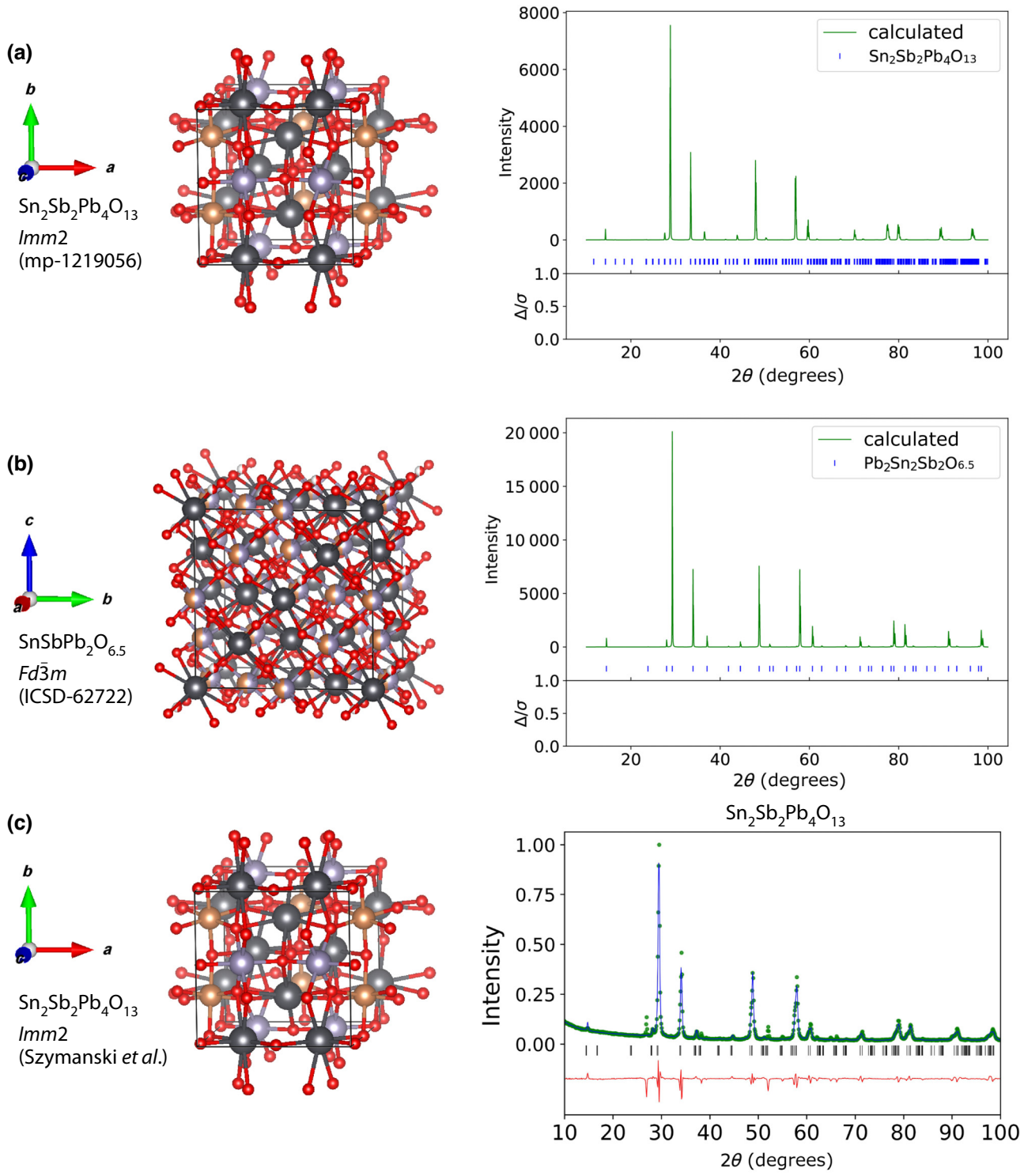

FIG. 3. (a) The structure of $Sn_2Sb_2Pb_4O_{13}$ as predicted by the Materials Project (mp-1219056, *Imm*2; left) and its simulated PXRD pattern (right). (b) The structure of the doped $SnSbPb_2O_{6.5}$ pyrochlore phase as listed in the ICSD (ICSD-62722, $Fd\bar{3}m$; left) and its simulated PXRD pattern (right). (c) The structure of $Sn_2Sb_2Pb_4O_{13}$ as provided by Szymanski *et al.* [4] (*Imm*2; left) and its reported PXRD pattern (right). Pb is shown in dark gray, Sn is shown in purple, Sb is shown in orange, and O is shown in red.

D. Spinels

Spinels, which possess an AB_2O_4 stoichiometry, are another common type of oxide material. A-lab has claimed to have synthesized the following spinel-derived compounds: $Zn_2Cr_3FeO_8$ ($R\bar{3}m$), $Zn_3Ni_4(SbO_6)_2$ ($C2/c$), and $Mn_4Zn_3(NiO_6)_2$ ($C2/c$). However, our refinements

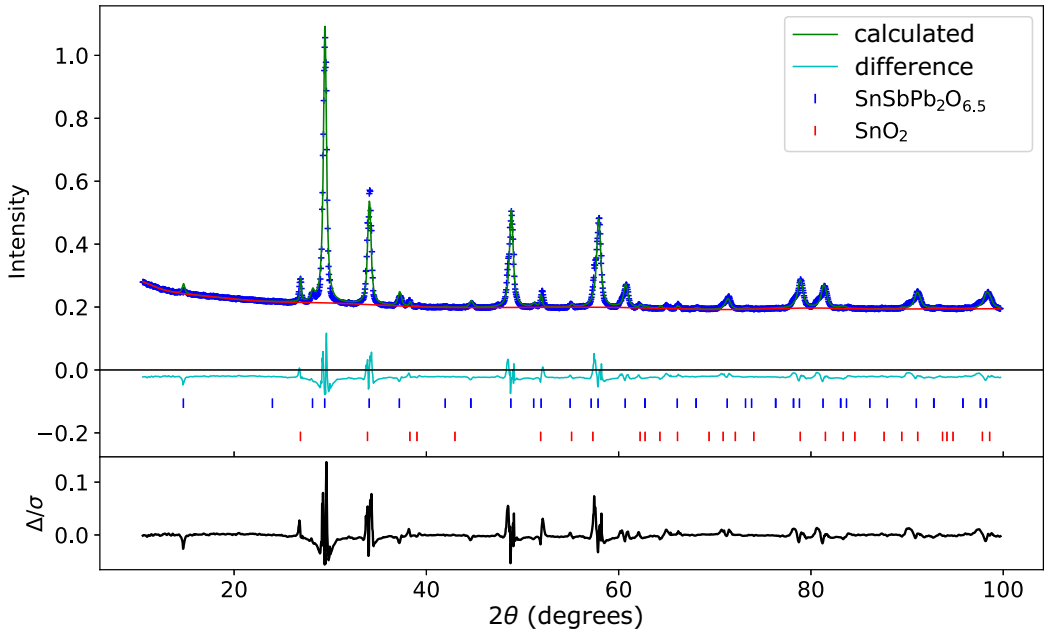

FIG. 4. Rietveld refinement of the experimental $Sn_2Sb_2Pb_4O_{13}$ pattern against the pyrochlore $SnSbPb_2O_{6.5}$ (coll-62722) and SnO_2 (coll-9163) phases as reported in the ICSD. Unlike the predicted $Sn_2Sb_2Pb_4O_{13}$ phase, which was supposed to crystallize in the orthorhombic space group Imm2, the pattern can be fitted well with the known higher-symmetry structure in space group $Fm\bar{3}m$.

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.

$NiMn_2O_4$ is a cubic spinel with complex cation and magnetic ordering. Guillemet-Fritsch *et al.* have explored Zn insertion into $NiMn_2O_4$ [29], including synthesis of a compound with almost exactly the same stoichiometry as the A-lab material $Mn_4Zn_3(NiO_6)_2$. In the work of Guillemet-Fritsch *et al.*, the A and B sites have been found to be compositionally disordered, with some Zn migration onto the octahedral sites, forming a disordered arrangement with the Mn and Ni ions (ICSD-92222–92224). However, the prediction from A-lab is of the A site (tetrahedral site) being 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 $Mn_4Zn_3(NiO_6)_2$ are refined against the existing compounds. The predominant phase in this pattern matches well with the disordered cubic spinel phase $(Zn_{0.759}Mn_{0.241})(Mn_{1.35}Ni_{0.65})O_4$ (ICSD-92223) in space group $Fd\bar{3}m$, with minority peaks indexing to NiO. The analysis of this compound, again, suffers from error 3.

Another synthesis to consider is the one that has targeted $Zn_3Ni_4(SbO_6)_2$. Gama *et al.* [30] have studied the inverse spinel $Zn_7(SbO_6)_2$ 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 $Zn_3Ni_4(SbO_6)_2$, i.e., the exact composition predicted in the A-lab paper. Gama *et al.* have 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 Supplemental Information. While the symmetry of both the structure predicted in the Materials Project and the structure reported in the A-lab paper have the same symmetry (space group $C2/c$), in the latter, the Ni and Sb ions are disordered (having fractional occupancies in the CIF) [4]. In Fig. 6, we compare the predicted structure from the Materials Project (mp-1216023) with its calculated diffraction pattern [see Fig. 6(a)], the reported ICSD structure from Gama *et al.* (ICSD-109468) with its calculated diffraction pattern [see Fig. 6(b)], and the structure provided in the Supplemental Information of Ref. [4] with the experimental pattern and A-lab fit [see Fig. 6(c)]. 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°, a feature that is much weaker in the disordered pattern. The experimental pattern clearly matches much better with the disordered version of $Zn_3Ni_4(SbO_6)_2$, which has been reported first by Gama *et al.* in 2003 (ICSD-109468).

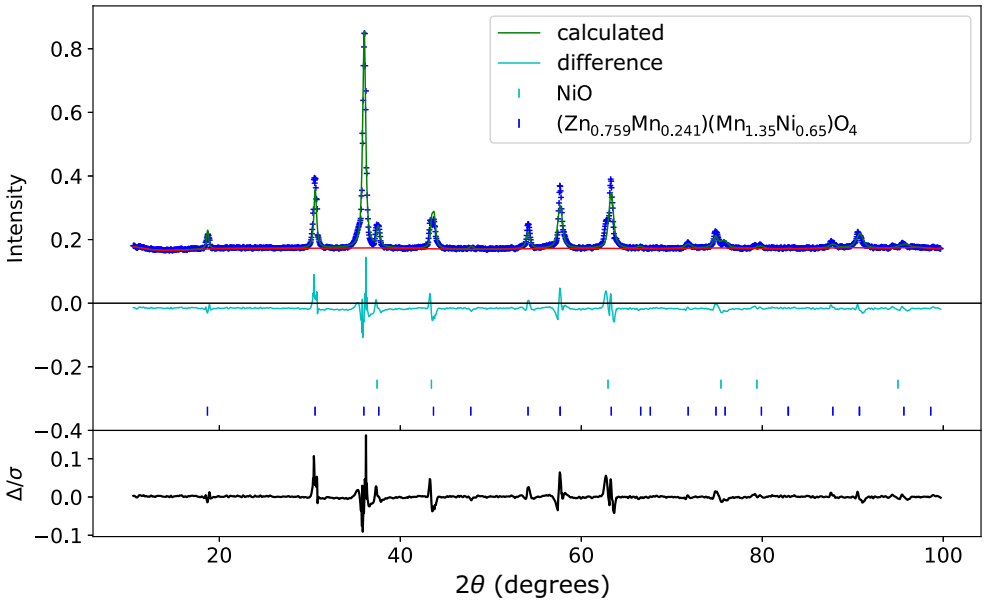

FIG. 5. Our Rietveld refinement of $Mn_4Zn_3(NiO_6)_2$ against the doped spinel phase $(Zn_{0.759}Mn_{0.241})(Mn_{1.35}Ni_{0.65})O_4$ (coll-92223) and NiO (coll-9866) as reported in the ICSD.

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 have 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) but note that other Ni contents give very similar fits and that 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° 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.

The compound $Zn_2Cr_3FeO_8$ 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 $ZnCr_2O_4$ and $ZnFe_2O_4$, including the exact composition predicted, were reported in 1970 and the B-site cations were described as disordered (ICSD-167363) [31]. Due to the small difference in scattering factor between Fe and Cr, there is only a minimal difference in the PXRD pattern between the predicted ordered structure and the known disordered structure. In addition, the reported PXRD pattern of $Zn_2Cr_3FeO_8$ also contains several unmodeled impurity peaks, which likely correspond to binary oxides of the metals. In the absence of any other evidence, we conclude that $Zn_2Cr_3FeO_8$ is not a new material and is the disordered B-site spinel described in 1970. Thus this falls under error 3.

E. Perovskites and ilmenites

Perovskite oxides have the general formula ABO_3 and 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 has reported the synthesis of six perovskite- or ilmenite-derived structures: Ba_2ZrSnO_6 ($Fm\bar{3}m$, perovskite), $MgTi_2NiO_6$ ($R3$, ilmenite), and $KBaGdWO_6$ 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 suggests 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 $Ba_6Na_2Ta_2V_2O_{17}$ [32]. There, the authors explain that both $Ba_6Na_2Ta_2V_2O_{17}$ and $Ba_6Na_2V_2Sb_{12}O_{17}$ have been mentioned as existing before but the structure has not been solved. They solve the structure as a 6C-type perovskite rather than the predicted 12H type, in disagreement with the materials prediction.

In the case of Ba_2SnZrO_6, the output of the refinement process, as published in the A-lab paper, contains many significant unmodeled diffraction peaks (Fig. 8).

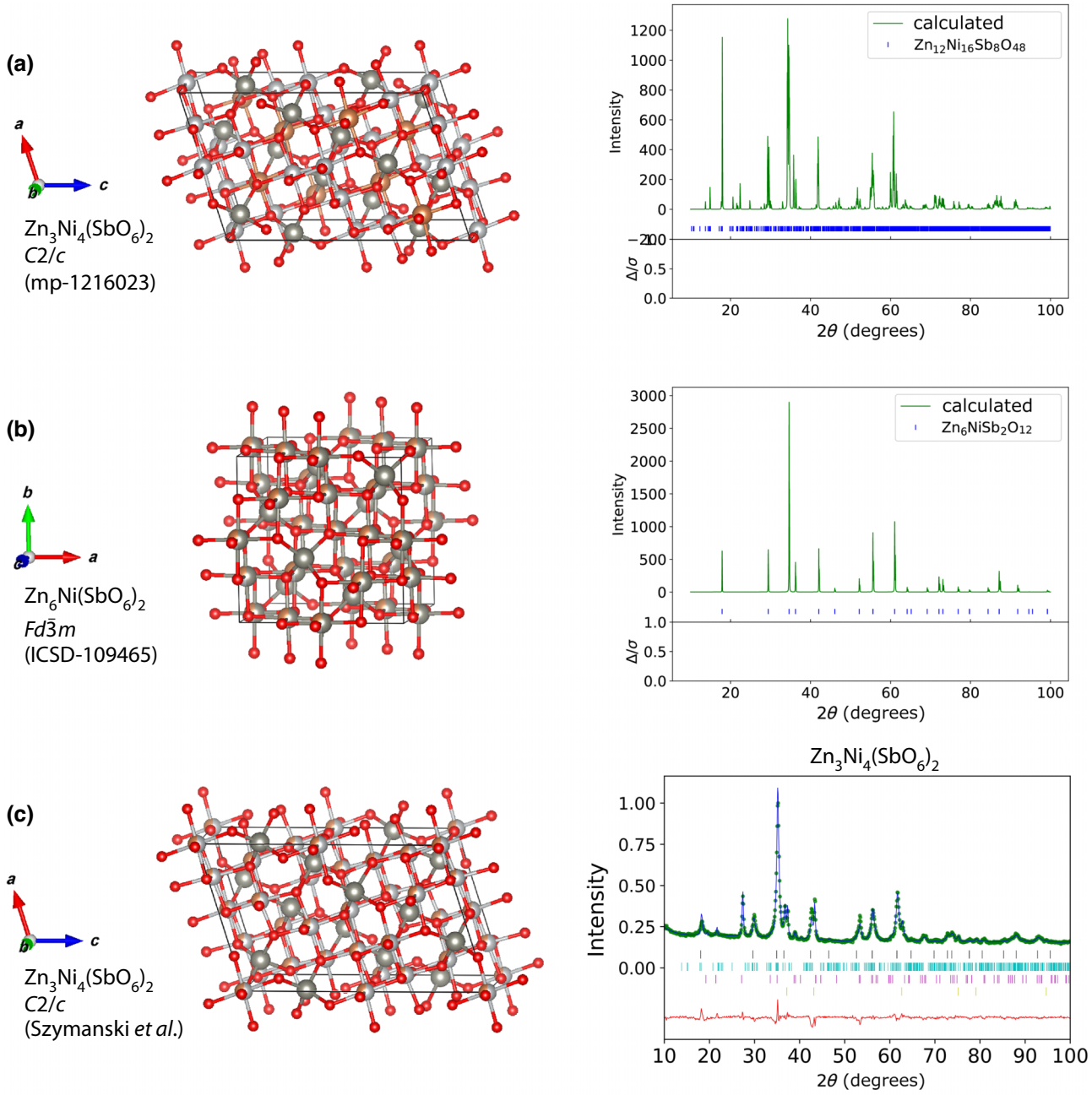

FIG. 6. (a) The structure of $Zn_3Ni_4(SbO_6)_2$ as predicted by the Materials Project (mp-1016023, $C2/c$; left) and its simulated PXRD pattern (right). (b) The structure of the doped $Zn_6NiSb_2O_{12}$ (coll-109465) phase as listed in the ICSD (ICSD-109465, $Fd\bar{3}m$; left) and its simulated PXRD pattern (right). (c) The structure of $Zn_3Ni_4(SbO_6)_2$ as provided by Szymanski *et al.* [4] ($C2/c$; left) and its reported PXRD pattern (right). Ni is shown in white, Zn is shown in gray, Sb is shown in orange, and O is shown in red.

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 unmodeled diffraction features. Our fit has 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 that 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

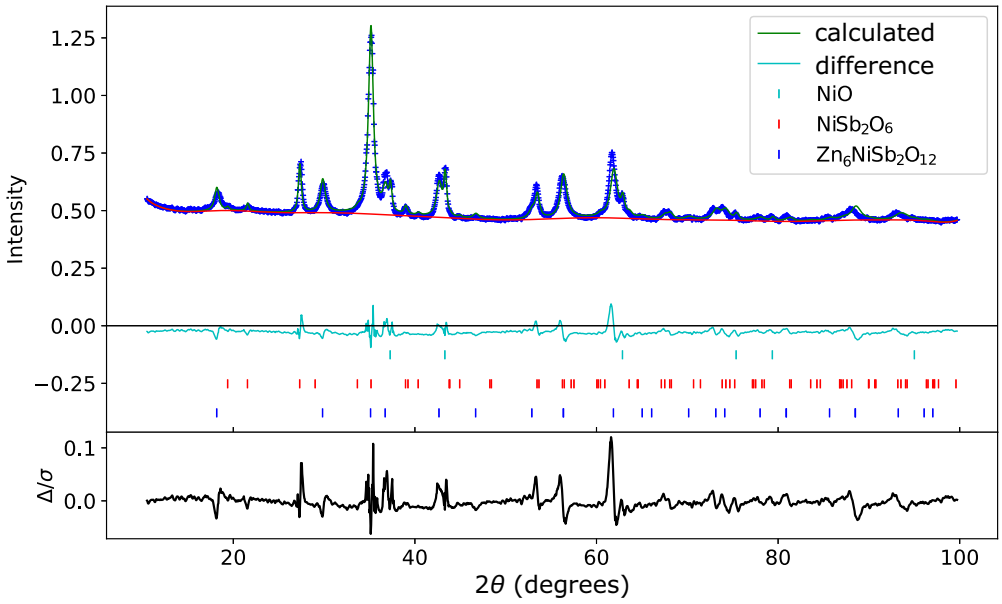

FIG. 7. (a) Rietveld refinement of the pattern representing $Zn_3Ni_4(SbO_6)_2$ against $Zn_6NiSb_2O_{12}$ (coll-109465), $NiSb_2O_6$ (coll-426852), and NiO (coll-9866), as reported in the ICSD.

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 thus suffers from errors 1 (poor fit), 2 (inconsistency in predicted and reported structure), and 3 (no evidence of cation order).

$MgTi_2NiO_6$ with space group $R3H$ is predicted to exist by A-lab as a new compound in a structure that is close to what is known in the 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 $R3H$, 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 between the ordered and disordered phases is very small. 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 [4], we find that the pattern for $MgTi_2NiO_6$ can be adequately fitted with a model of the known ilmenite phase $Ni_{0.5}Mg_{0.5}TiO_3$ (ICSD-171583) and a small impurity (2 wt%) of NiO. This, again, is an example of error 3. 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 [4]. 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 that 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.

For the double perovskites $KBaGdWO_6$ and $KBaPrWO_6$, 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 [33]—a classic example is the cuprate superconductor $YBa_2Cu_3O_7$. In the absence of oxygen deficiency, A-site order can rarely occur but has, to our knowledge, not been seen in the

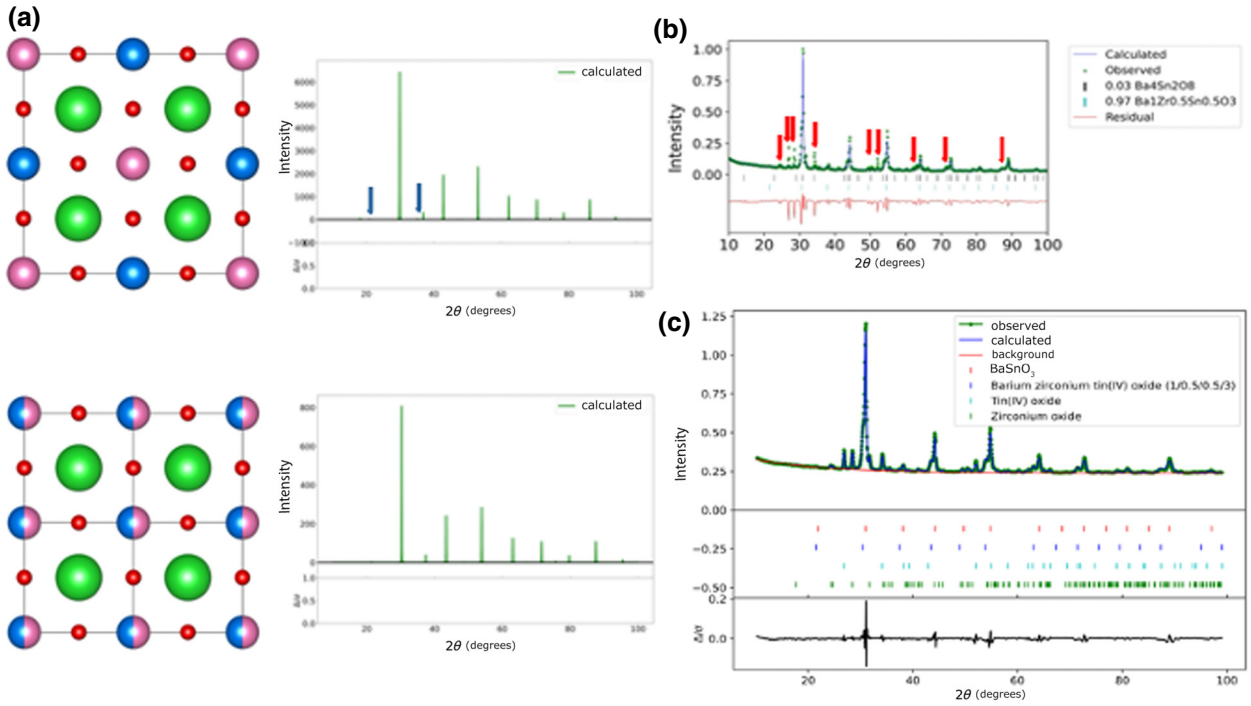

FIG. 8. (a) A comparison of the predicted structure of Ba_2ZrSnO_6 (mp-1228067, $Fm\bar{3}m$; upper left—ICSD-43137, $Pm\bar{3}m$; lower left) with the structure of known $BaSn_{0.5}Zr_{0.5}O_3$ (right). The simulated powder patterns of both phases are very similar. The blue arrows show the positions of the two most intense peaks that appear in the pattern from the ordered structure but not that of the disordered structure. (b) The refinement of Ba_2ZrSnO_6 as given in Ref. [4], 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.

symmetry predicted here—typically, A sites form ordered layers driven by large charge differences, rather than fcc patterns [34]. 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 characterize the ordering. This evidence is not supplied by the poor fits to the PXRD data.

F. Phosphates

Phosphates are ionic compounds that contain 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 $P_2O_7^{4-}$ anions), or metaphosphate- (PO_3^-) containing phases, given as $K_2TiCr(PO_4)_3$, $CaFe_2P_2O_9$, $CaCo(PO_3)_4$, $CaMn(PO_3)_4$, $CaNi(PO_3)_4$, $InSb_3(PO_4)_6$, $K_4MgFe_3(PO_4)_5$, $K_4TiSn_3(PO_5)_4$, $KNaTi_2(PO_5)_2$, $KNaP_6(PbO_3)_8$, $MgNi(PO_3)_4$, $MgTi_4(PO_4)_6$, $Na_3Ca_{18}Fe(PO_4)_{14}$, $Na_7Mg_7Fe_5(PO_4)_{12}$, $NaMnFe(PO_4)_2$, Mn_2VPO_7, $Mn_7(P_2O_7)_4$, and $MgCuP_2O_7$.

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 1 (poor fit) and 3 (no evidence of cation order) are very common in the analysis of the phosphates.

The compound $K_2TiCr(PO_4)_3$ has been predicted to exist as a new cubic phase in the space group $P2_13$. Figure 10(d) shows our refinement of the provided PXRD pattern, using known cubic $K_2Ti_2(PO_4)_3$ ($P2_13$; ICSD-202888) and Cr_2O_3, a common impurity in high-temperature synthesis of oxides containing chromium [35]. The refinement provided in the A-lab paper has several unfitted peaks, all of which correspond to the Cr_2O_3 impurity phase as marked by the red arrows in Fig. 10(c). The example of $K_2TiCr(PO_4)_3$ 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 III). We consider it to be the responsibility of the authors of the A-lab paper [4] to unambiguously prove the synthesis of the target materials in all cases and we will refrain from providing alternative refinements of all 43 materials in this

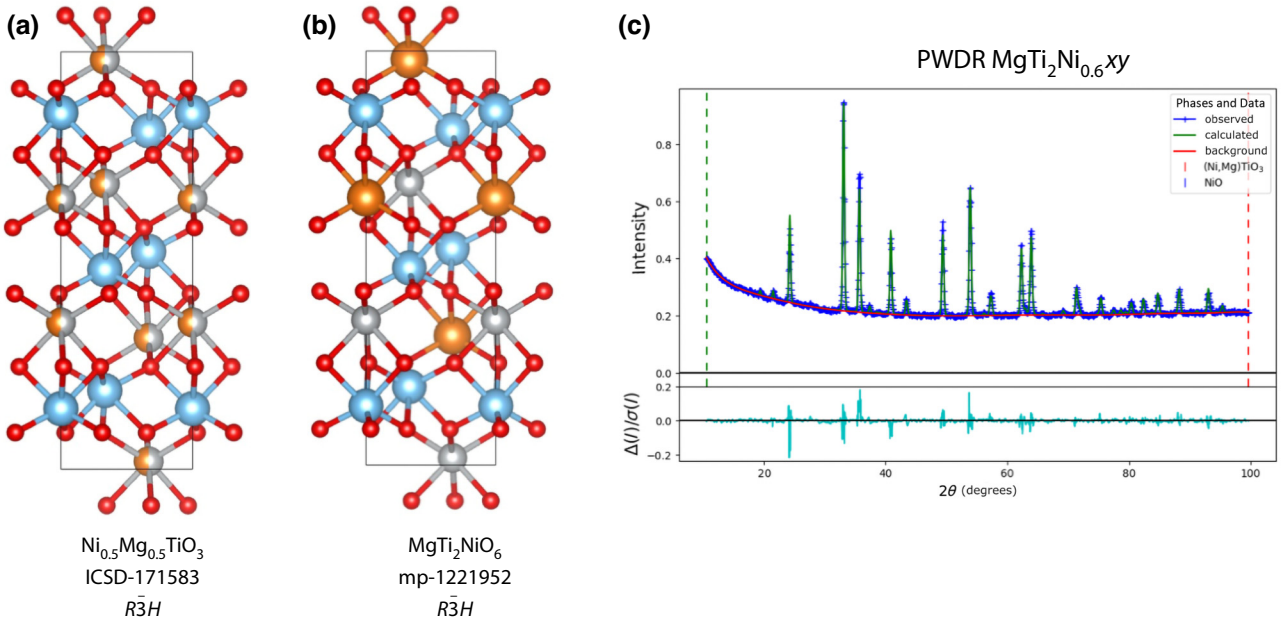

FIG. 9. (a),(b) A comparison of (a) the predicted structure of $MgTi_2NiO_6$ (ICSD-171583, $R\bar{3}H$) with (b) the structure of known $Ni_{0.5}Mg_{0.5}TiO_3$ (mp-1221952, $R3H$). (c) Refinement of the corresponding PXRD pattern using $Ni_{0.5}Mg_{0.5}TiO_3$, yielding a satisfactory fit.

Perspective. We will, however, discuss each compound and possible alternatives briefly below.

$KNaP_6(PbO_3)_8$ has been 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 [4] has a higher symmetry and the space group $P6_3/m$, although the authors incorrectly claim they have synthesized the target compound in space group $P3$. The structure provided in the Supplemental Information not only has a higher symmetry but it also places K, Na, and Pb on the same crystallographic position. We have found that the PXRD pattern of the synthetic product could also be indexed with a combination of the known materials $(Na,K)Pb_4(PO_4)_3$ and $Pb_8O_5(PO_4)_2$ (ICSD-182501, 98702), which provides an alternative interpretation of the data.

There are four predicted compounds in the A-lab data set based on the $Ni_2(PO_3)_4$ structure, 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, and 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 synthesized previously, namely, (M,M') = (Ni,Zn), (Ni,Co), and (Mg,Mn). Nord has investigated the ordering for bimetallic tetrametaphosphates with (M,M') = (Ni,Zn) and (Ni,Co), using neutron diffraction, as XRD has only small scattering-factor differences between these metals [36]. It has been 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 has been observed [37]. Experimental data are presented by 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 $(M,M')_2(PO_3)_4Ca$ has entered the tetrametaphosphate phase, as the PXRD patterns 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 the use of alternative diffraction techniques such as neutron

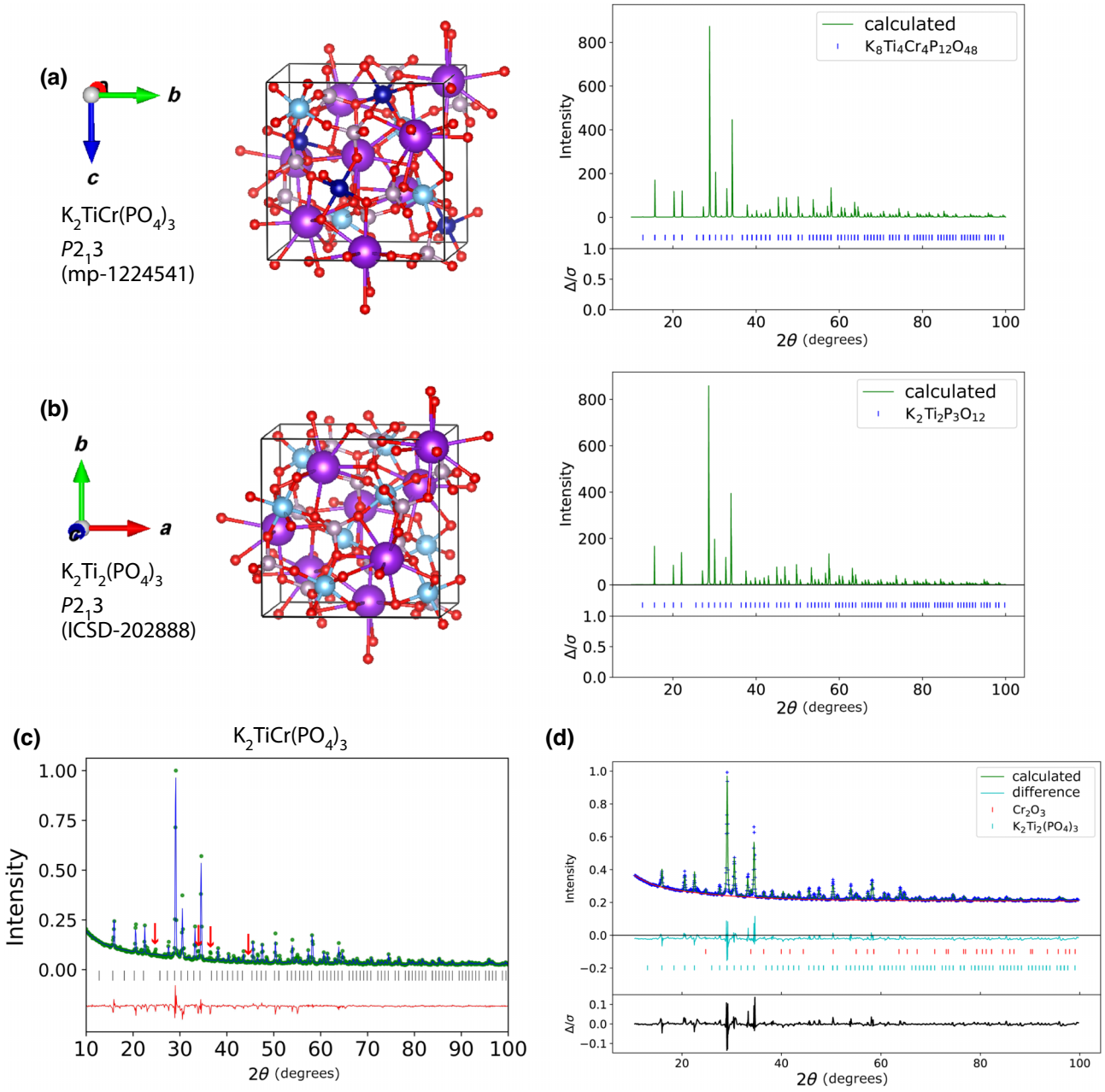

FIG. 10. (a) The structure of $K_2TiCr(PO_4)_3$ as predicted by the Materials Project (mp-1224541, $P2_13$; left) and its simulated PXRD pattern (right). (b) The structure of the $K_2Ti_2(PO_4)_3$ phase as listed in the ICSD (ICSD-202888, $P2_13$; left) and its simulated PXRD pattern (right). (c) The reported PXRD pattern of $K_2TiCr(PO_4)_3$ as provided by Szymanski *et al.* [4], 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 in the ICSD. The unmatched peaks in the refinement provided in the A-lab paper are matched to the Cr_2O_3 phase. K is shown in purple, Ti is shown in sky blue, Cr is shown in navy blue, P is shown in light purple, and O is shown in red.

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 for the others discussed in this paragraph.

The compound $NaMnFe(PO_4)_2$ 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, and Fe metals in different proportions, which are included in the ICSD [38]. We consider this the more likely identity of the material.

The compound Mn_2VPO_7 is the thortveitite structure, which is known for both $Mn_2P_2O_7$ and $Mn_2V_2O_7$. 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 $Mn_2P_2O_7$ or $Mn_2V_2O_7$ by PXRD requires careful measurement of the peak intensities, which change only by a small amount, as V is reasonably close to P in electron density. This analysis has not been 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 $MgTi_4(PO_4)_6$ is a known compound, reported by Barth *et al.* in 1993 (ICSD-74287) [39]. The proposed structure has ordered Mg-ion vacancies. If the Mg-ion vacancies were disordered, then the small peak seen experimentally at around 17° would be absent. This peak has also been observed by Barth *et al.*, who have 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 [40].

The predicted compound $KNaTi_2(PO_5)_2$ is a cation-ordered version of a series of titanyl phosphates that have been studied by various groups; e.g., Norberg *et al.* have reported one such disordered version with a very similar stoichiometry to that predicted here in 2003 [41] and an earlier report is listed in the ICSD (ICSD-71239). 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 have not been done, so there is no evidence of the ordered phase. It is likely that the previously reported disordered phase has been made.

The predicted compound $MgCuP_2O_7$ is based on mp-1041741. However, the CIF file provided in the Supplemental Information of the A-lab paper 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, which is not in the ICSD but has been reported in 1990 [42].

To complete this section, we feel that the fits to the models for the predicted compounds $Na_3Ca_{18}Fe(PO_4)_{14}$, $Na_7Mg_7Fe_5(PO_4)_{12}$, $Mn_7(P_2O_7)_4$, and $K_4TiSn_3(PO_5)_4$ 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.

G. 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 $MgV_4Cu_3O_{14}$ ($P1$), $CaGd_2Zr(GaO_3)_4$ ($P1$), $Ba_9Ca_3La_4(Fe_4O_{15})_2$ ($P1$), $KNa_2Ga_3(SiO_4)_3$ ($P2_1/c$),$KPr_9(Si_3O_{13})_2$ ($P3$), and $NaCaMgFe(SiO_3)_4$ ($C2$). With the exception of $KPr_9(Si_3O_{13})_2$, 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 number of fitting parameters makes refinement extremely challenging and a lot of care needs to be taken to ensure that the low symmetry is real. The data provided do not satisfactorily prove the existence of these phases and for all of the low-symmetry phases above, we have found alternative matches to the PXRD data, which could explain the synthetic product to be a combination of known materials. For example, $CaGd_2Zr(GaO_3)_4$ is likely a type of garnet, e.g., cubic $Ca_{0.95}Zr_{0.95}Gd_{2.05}Ga_{4.05}O_{12}$ (ICSD-202850, space group $Ia\bar{3}d$), which matches the main peaks of the pattern well. $NaCaMgFe(SiO_3)_4$ is predicted based on mp-1221075. However, the Materials Project entry has ordered cations, while the CIF provided in the Supplemental Information of the A-lab paper 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 is listed in the ICSD (e.g., ICSD-417169). Lenaz *et al.* has described one study into these compounds, which have two crystallographic cation sites, one of which has disordered Ca and Na, while the other has disordered Mg and Fe [43]. The experimental pattern from A-lab matches well with the known compounds.

$KPr_9(Si_3O_{13})_2$, which is predicted to have higher symmetry ($R3$), again possesses disorder in the CIF file provided in the A-lab paper [4] (between K and Pr), which has not been predicted in the original structure. It thus follows a similar theme to many compounds that we have already discussed in detail. $MgV_4Cu_3O_{14}$ is the same composition as an existing compound in the ICSD, $Cu_{1.5}Mg_{0.5}V_2O_7$ (ICSD-69731). The predicted compound has ordered Mg and Cu ions and is based on the $Cu_2P_2O_7$ 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°, which is absent in the disordered pattern. Sadly, the A-lab data do not go below 10°, so this confirmatory peak has not been measured. There is no evidence from the PXRD data that the ordered compound has been made over the disordered.

H. 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 $MnAgO_2$, $Y_3In_2Ga_3O_{12}$, and $CaFe_2(PO_4)_2O$, which have been reported, respectively, in Refs. [44–46]. Of those $CaFe_2(PO_4)_2O$ seems to have been convincingly synthesized based on the provided PXRD data, whereas the PXRD patterns for the other two are fitted so poorly that it is difficult to state whether the materials have indeed 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 have been reported relatively recently, between 2021 and 2023. In fact, the authors of the Google DeepMind paper [3] have 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 have since been discovered.

V. OVERVIEW

Above, we have outlined issues with the "new" materials synthesized by A-lab. Of the 36 compounds classed as "successes" and the seven classified as "partial successes" by A-lab (43 in total), we have found significant issues with 42 of them (the exception being $CaFe_2(PO_4)_2O$). 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, $MnAgO_2$ and $Y_3In_2Ga_3O_{12}$. In this case, the success rate would be 3/58, or 5%, which is far away from the claims in the paper. A stricter interpretation of materials discovery (but one usually applied to traditional laboratories claiming a new discovery), that a new material must not have been published previously at all, leads to the conclusion that none of the materials in the A-lab paper would be counted as discoveries.

We have 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 have found might also point to a general issue with the materials-prediction part of the work, namely, they can often be derived from known compounds in which cation order breaks symmetry. As density-functional theory (DFT) cannot model compositionally disordered atoms easily, there might be an underlying error in the way in which 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 these 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" [47]. We hope that our Perspective has 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 40/43 cases, be interpreted with the opposite conclusion, which is that in these cases, the predicted materials have not been successfully synthesized. Even if those known materials were not in the training set and thus were "new to the A-lab," as claimed by the authors in response, this makes the tile and abstract, which claim "novel materials," highly misleading. Major improvements to the data analysis and the addition of careful compositional characterization are necessary to draw the alleged conclusions.

VI. 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 realize that no value of R_{wp} or χ^2 alone can ever be sufficient to conclude that a model is even approximately correct. Toby has 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" [48]. What role does this "human intuition" in assessing the quality of a fit play and how can this be replicated by machine-learning models? To begin, we can consider why goodness-of-fit statistics alone are insufficient. Not all goodness of fit is equal: in a Rietveld refinement, completely unfitted diffraction features may increase the χ^2 value 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; e.g., that the real material has a different symmetry to the one modeled.

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 fit 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 above discussion, 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 $SrTiO_3$ 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 Sr_2FeMoO_6, where the Fe and Mo B site cations can show different degrees of disorder [49]. 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.* have matched experimental patterns of organic polymorphs to a database of calculated structures [50]. Lunt *et al.* have extended this methodology to identify polymorphs of organic materials crystallized and analyzed in a robotic lab [51]. Salgardo *et al.* have used machine learning to classify crystal system and space group from experimental PXRD patterns [52]. 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.

In general, powder diffraction is not necessarily the best method to solve a new crystal structure. Single-crystal x-ray diffraction is much more accurate in this endeavor but, of course, requires the synthesis of (small) single crystals of the product. Maybe the rate of new-materials discovery can be increased if the analysis step is replaced with single-crystal x-ray diffraction, as such data are often more reliable in providing unambiguous information of the structure of the products.

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 recognized. An early approach to the modeling of 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 [53]. While the VCA has been used to calculate some compositionally disordered oxides, it is generally recognized as having important limitations [54]. The bonding properties of compositionally disordered materials often cannot be successfully modeled by averaged virtual atoms and this can lead to very large divergence between VCA and experiment [55]. 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 [56]. In different ways, both VCM and CPA 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 of this type currently in use. The cluster-expansion (CE) model began

with work to use the Ising model of magnetism to describe compositional order or disorder in the 1950s [57] and has seen major advances from the 1980s [58] onward. CE considers finite-size clusters and computes the properties of the material as a combination of these. Zunger *et al.* have introduced the special quasirandom structure (SRS) as a way of approximating random distributions of atoms within a finite supercell [59]. Grau-Crespo *et al.* have used the concept of calculating all possible configurations in a given cell size that represent the total composition of a disordered material [60]. 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 [61]. This very brief survey is provided to show that the issues presented by modeling disordered materials are not new. To our knowledge, the Materials Project does not use any of the approaches that 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.

It might be argued that compositional order does not matter and that a prediction of an ordered compound is really the same as that of a disordered one. We reject such an argument. Ordered and disordered compounds have such important differences in structure and properties that they cannot be considered the same. In oxides, this is especially well studied. The degree of spinel inversion (a type of compositional disorder) dictates magnetic and transport properties [62]. In perovskites, disorder influences catalytic, magnetic, transport, and ferroelectric properties [63]. It is untenable to conflate disordered with ordered compounds and assume that they are the same but worse to claim a prediction of an ordered compound is new when experiment produces only the disordered analogue, which is already known.

That two thirds of A-lab's "successful" syntheses are likely unrecognized 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 at all is made of compositional disorder.

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 the 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 the ICSD) but some may take issue with this definition.
(2) When predicting new nonmolecular inorganic materials computationally, the possibility of compositional disorder should be considered explicitly.
(3) When analyzing 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 [64]: positive hypothesis testing should not be used in materials discovery and 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 toward unsupervised analysis have been made but in our view, truly autonomous material analysis remains a target for future work, or a better developed human-machine interface might drastically help to speed up the process. We hope that this Perspective outlines some of the pitfalls and will help to strengthen this aspect in the future.

ACKNOWLEDGMENTS

L.M.S. was supported by the National Science Foundation (NSF) through NSF Grant No. OAC-2118310 and the through the Princeton Center for Complex Materials, an NSF Materials Research Science and Engineering Center (MRSEC), Grant No. DMR-2011750 as well as by the Gordon and Betty Moore Foundation's Emergent Phenomena in Quantum Systems (EPIQS) initiative through Grant No. GBMF9064 and the David and Lucille Packard foundation. S.B.L. is supported by the NSF 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 NSF. Y.L. was supported by the China Scholarship Council.

APPENDIX

1. Indexing of new phases

TABLE III. The table of new compounds synthesized by A-lab compared against potential matches found in the ICSD. 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 [65] 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 the 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_2ZrSnO_6	$Fm\bar{3}m$ (225)	$Pm\bar{3}m$ (221)	$BaSnO_3$ $Ba(Zr_{0.5}Sn_{0.5}O_3)$ SnO_2 ZrO_2	188149 43137 9163 66781	$Pm\bar{3}m$ (221) $Pm\bar{3}m$ (221) $P4_2/mnm$ (136) $P4_2/nmc$ (137)
$Ba_6Na_2Ta_2V_2O_{17}$	$P6_3mmc$ (194)	$P6_3mmc$ (194)	$BaNaVO_4$ $NaVO_3$	130002 2103	$P\bar{3}m1$ (164) C12/c1 (15)
$Ba_6Na_2V_2Sb_2O_{17}$	$P6_3mmc$ (194)	$P6_3mmc$ (194)	$BaNaVO_4$ $NaVO_3$	130002 2103	$P\bar{3}m1$ (164) C12/c1 (15)
$Ba_9Ca_3La_4(Fe_4O_{15})_2$	P1 (1)	P1 (1)	$Ba_{4.5}Ca_{1.5}La_2Fe_4O_{15}$	72336	$P6_3mc$ (186)
$CaCo(PO_3)_4$	C2/c (15)	C2/c (15)	$Co_2P_2O_7$ $Ca_3(PO_4)_2$	2830 923	$B12_1/c1$ (14) $P12_1/a1$ (14)
$CaFe_2P_2O_9$	Pnma (62)	Pnma (62)	No matches	-	-
$CaGd_2Zr(GaO_3)_4$	$P\bar{1}$ (2)	$P\bar{1}$ (2)	$Ca_{0.95}Zr_{0.95}$— $Gd_{2.05}Ga_{4.05}O_{12}$	202850	$Ia\bar{3}d$ (230)
$CaMn(PO_3)_4$	C2/c (15)	C2/c (15)	$Mn_2P_4O_{12}$	412558	C12/c1 (15)
$CaNi(PO_3)_4$	C2/c (15)	C2/c (15)	$CaNi_3(P_2O_7)_2$ $Ca_2(P_2O_7)$	74046 14313	$P12_1/c1$ (14) $P4_1$ (76)
$FeSb_3Pb_4O_{13}$	R3m (160)	R3m (160)	$Pb_2(Fe,Sb)O_{6.5}$	60805	$Fd\bar{3}m$ (227)
$Hf_2Sb_2Pb_4O_{13}$	Imm2 (44)	Imm2 (44)	$Pb_2(Hf,Sb)O_{6.5}$ $PbHfO_3$	62723 174110	$Fd\bar{3}m$ (227) Pbam (55)
$InSb_3(PO_4)_6$	Pc (7)	Pc (7)	$Sb_{0.5}In_{0.5}(P_2O_7)$ $Sb(Sb_{0.5}In_{0.5})(PO_4)_3$	166834 166835	$P12_1/n1$ (14) $Pna2_1$ (33)
$InSb_3Pb_4O_{13}$	R3m (160)	R3m (160)	$Pb_2(InSb)O_{6.5}$ In_2O_3	41119 14388	$Fd\bar{3}m$ (227) $Ia\bar{3}$ (206)
$K_2TiCr(PO_4)_3$	$P2_13$ (198)	$P2_13$ (198)	$K_2Ti_2(PO_4)_3$ Cr_2O_3	202888 626479	$P2_13$ (198) $R\bar{3}c$ (167)
$K_4MgFe_3(PO_4)_5$	Cc (9)	Cc (9)	$K_4MgFe_3(PO_4)_5$	161484	$P\bar{4}2_1c$ (114)
$K_4TiSn_3(PO_5)_4$	$P2_1$ (4)	$P2_1$ (4)	$K((Ti_{0.25}Sn_{0.75})O)(PO_4)$	250088	$Pna2_1$ (33)
$KBaGdWO_6$	$F\bar{4}3m$ (216)	$F\bar{4}3m$ (216)	Ba_2GdWO_6 Gd_2O_3	138973 40473	$Fm\bar{3}m$ (225) $Ia\bar{3}$ (206)
$KBaPrWO_6$	$F\bar{4}3m$ (216)	$F\bar{4}3m$ (216)	$Ba_{11}W_4O_{23}$ Pr_6O_{11} (EVA)	418207 N/A	$Fd\bar{3}m$ (227) $P2_1/c$ (14)
KMn_3O_6	C2/c (15)	C2/m (12)	$K_{0.48}Mn_{1.94}O_{5.18}$	240249	$P6_3/mmc$ (194)
$KNa_2Ga_3(SiO_4)_3$	$P2_1/c$ (14)	$P2_1/c$ (14)	$NaGaSiO_4$	46861	$P12_1/n1$ (14)

TABLE III. Continued

Proposed Phase	Proposed Symmetry	Indexed Symmetry	ICSD Phase	ICSD Code	ICSD Symmetry
$KNaP_6(PbO_3)_8$	P3 (143)	$P6_3/m$ (176)	$(Na,K)Pb_4(PO_4)_3$ $Pb_8O_5(PO_4)_2$	182501 98702	$P6_3/m$ (176) C12/m1 (12)
$KNaTi_2(PO_5)_2$	$Pna2_1$ (33)	$Pna2_1$ (33)	$Na_{0.5}(Na_{0.492}K_{.008})$— $(TiO)(PO_4)$	59284	$Pna2_1$ (33)
$KPr_9(Si_3O_{13})_2$	P3 (143)	P3 (143)	$Pr_{9.33}Si_6O_{32}$ (EVA)	N/A	-
$Mg_3MnNi_3O_8$	$R\bar{3}m$ (166)	$R\bar{3}m$ (166)	NiO $(Ni,Mn)(Ni,Mn)_2O_4$	9866 84517	$Fm\bar{3}m$ (225) $Fd\bar{3}m$ (227)
Mg_3NiO_4	$Pm\bar{3}m$ (221)	$Fm\bar{3}m$ (225)	$MgNiO_2$	290603	$Fm\bar{3}m$ (225)
$MgCuP_2O_7$	$P\bar{1}$ (2)	$P\bar{1}$ (2)	$Mg_2P_2O_7$	20295	C12/m1 (12)
$MgNi(PO_3)_4$	C2/c (15)	C2/c (15)	$Mg_3(PO_4)_2$ $Ni_3(PO_4)_2$ $(Ni,Mn)_3(PO_4)_2$	31005 153159 158525	$P12_1/n1$ (14) $P12_1/c1$ (14) $P12_1/a1$ (14)
$MgTi_2NiO_6$	R3 (146)	R3 (146)	$(Ni_{0.5}Mg_{0.5})TiO_3$ TiO	171583 38755	$R\bar{3}$ (148) $Fm\bar{3}m$ (225)
$MgTi_4(PO_4)_6$	R3 (146)	R3 (146)	$Mg_{0.5}Ti_2(PO_4)_3$	74287	$R\bar{3}c$ (167)
$MgV_4Cu_3O_{14}$	P1 (1)	P1 (1)	$(Cu_{1.5}Mg_{0.5})V_2O_7$	69731	C12/c1 (15)
Mn_2VPO_7	Cm (8)	Cm (8)	$Mg_2P_2O_7$	47136	C12/m1 (12)
$Mn_4Zn_3(NiO_6)_2$	C2/c (15)	C2/c (15)	$(Zn_{0.759}Mn_{0.241})$— $(Mn_{1.35}Ni_{0.65})O_4$ NiO	92223 9866	$Fd\bar{3}m$ (227) $Fm\bar{3}m$ (225)
$Mn_7(P_2O_7)_4$	$C222_1$ (20)	$C222_1$ (20)	$Mn_2(PO_3)_4$	145534	C12/c1 (15)
$MnAgO_2$	C2/m (12)	C2/m (12)	$MnAgO_2$	139006	C12/m1 (12)
$Na_3Ca_{18}Fe(PO_4)_{14}$	P1 (1)	P1 (1)	$Na_3Ca_{18}Fe(PO_4)_{14}$	85103	R3c (161)
$Na_7Mg_7Fe_5(PO_4)_{12}$	P1 (1)	P1 (1)	$Na_2Mg_2Fe(PO_4)_3$	138263	C12/c1 (15)
$NaCaMgFe(SiO_3)_4$	C2 (5)	C2 (5)	$(Ca_{0.774}Na_{0.226})$— $(Mg_{0.901}Fe_{0.099})$— $Fe_{0.011}(Si_2O_6)$	75294	C12/c1 (15)
$NaMnFe(PO_4)_2$	$P\bar{1}$ (2)	$P\bar{1}$ (2)	$Na_2Mg_2Fe(PO_4)_3$	138263	C12/c1 (15)
$Sn_2Sb_2Pb_4O_{13}$	Imm2 (44)	Imm2 (44)	$Pb_2SnSbO_{6.5}$	62722	$Fd\bar{3}m$ (227)
$Y_3In_2Ga_3O_{12}$	$Ia\bar{3}d$ (230)	$Ia\bar{3}d$ (230)	$Y_3In_2Ga_3O_{12}$	54664	$Ia\bar{3}d$ (230)
$Zn_2Cr_3FeO_8$	$R\bar{3}m$ (166)	$R\bar{3}m$ (166)	$(Zn_{0.54}Fe_{0.46})Fe_2O_4$ $Zn(FeCrO_4)$	81207 167362	$Fd\bar{3}m$ (227) $Fd\bar{3}m$ (227)
$Zn_3Ni_4(SbO_6)_2$	C2/c (15)	C2/c (15)	$Zn(Zn_{1.333}Sb_{0.667})O_4$ NiO $NiSb_2O_6$	173996 9866 426852	$Fd\bar{3}m$ (227) $Fm\bar{3}m$ (225) $P4_2/mnm$ (136)
$Zr_2Sb_2Pb_4O_{13}$	Imm2 (44)	Imm2 (44)	$Pb_2(ZrSb)O_{6.5}$	62721	$Fd\bar{3}m$ (227)

2. Additional refinements

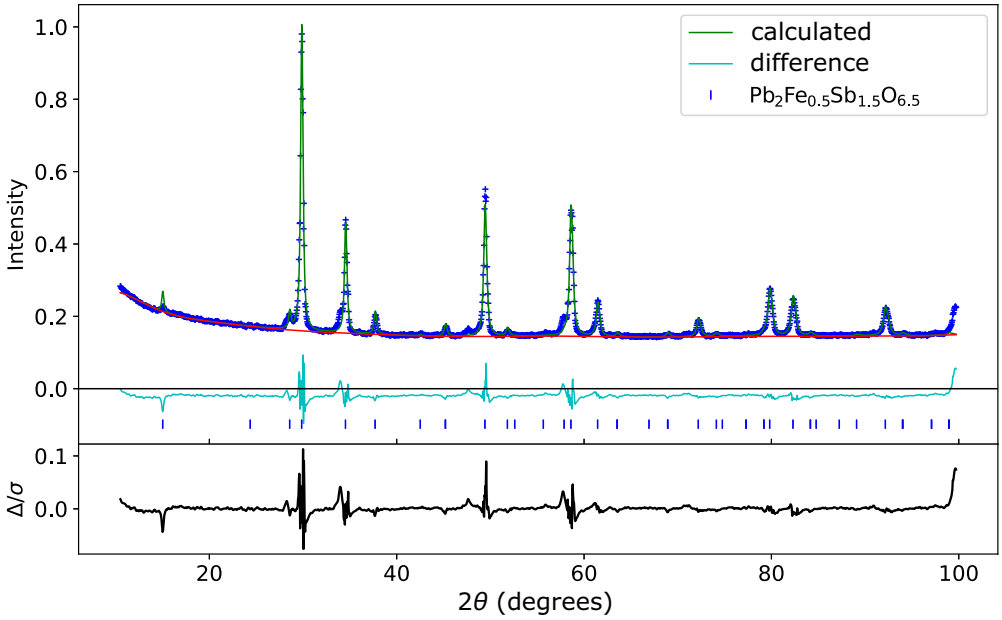

FIG. 11. Rietveld refinement of the pattern representing $FeSb_3Pb_4O_{14}$ against $Pb_2Fe_{0.5}Sb_{1.5}O_{6.5}$ (coll-60805), as reported in the ICSD.

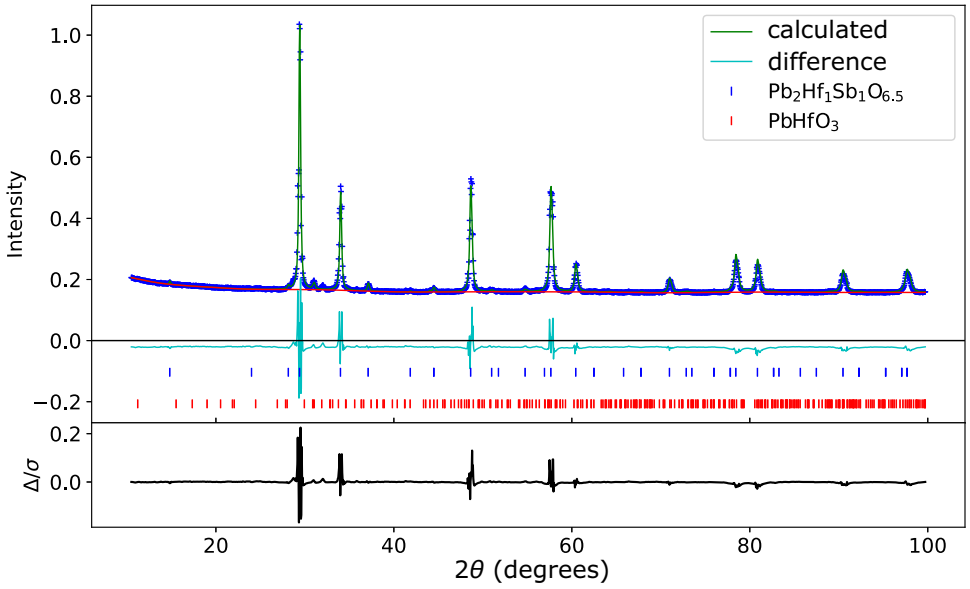

FIG. 12. Rietveld refinement of the pattern representing $Hf_2Sb_2Pb_4O_{13}$ against $Pb_2HfSbO_{6.5}$ (coll-60805) and $PbHfO_3$ (coll-174110), as reported in the ICSD.

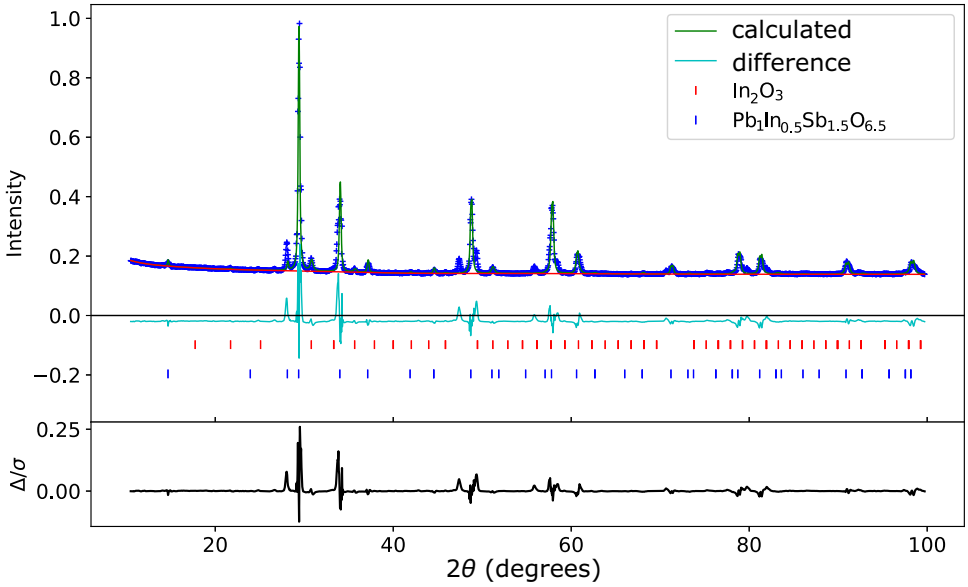

FIG. 13. Rietveld refinement of the pattern representing $InSb_3Pb_4O_{13}$ against $Pb_1In_{0.5}$ $Sb_{1.5}$ $O_{6.5}$ (coll-14388) and In_2O_3 (coll-41119), as reported in the ICSD.

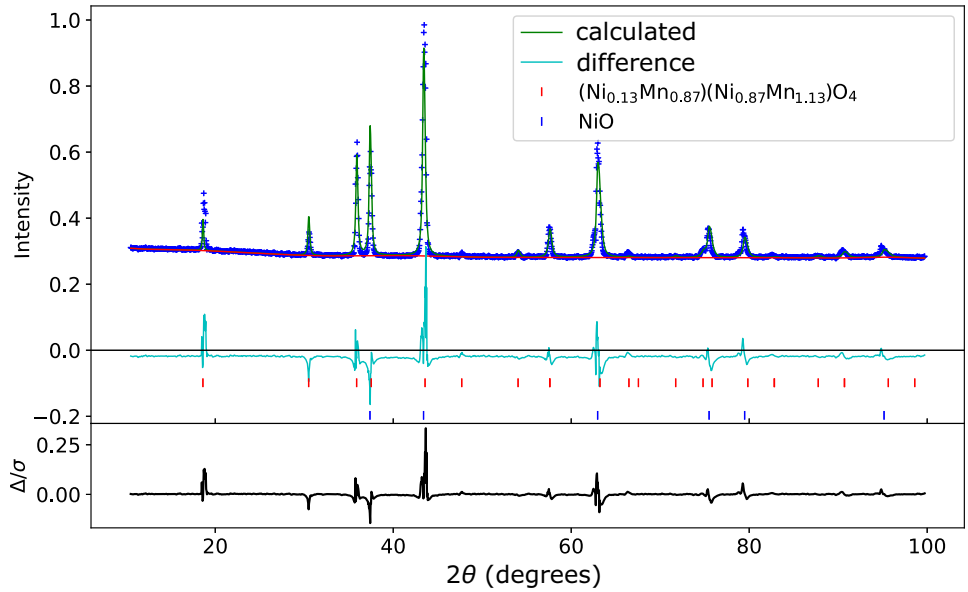

FIG. 14. Rietveld refinement of the pattern representing $Mg_3MnNi_3O_8$ against $(Ni_{0.13}$ $Mn_{0.87})$ $(Ni_{0.87}$ $Mn_{1.13})$ O_4 (coll-84517) and NiO (coll-9866), as reported in the ICSD.

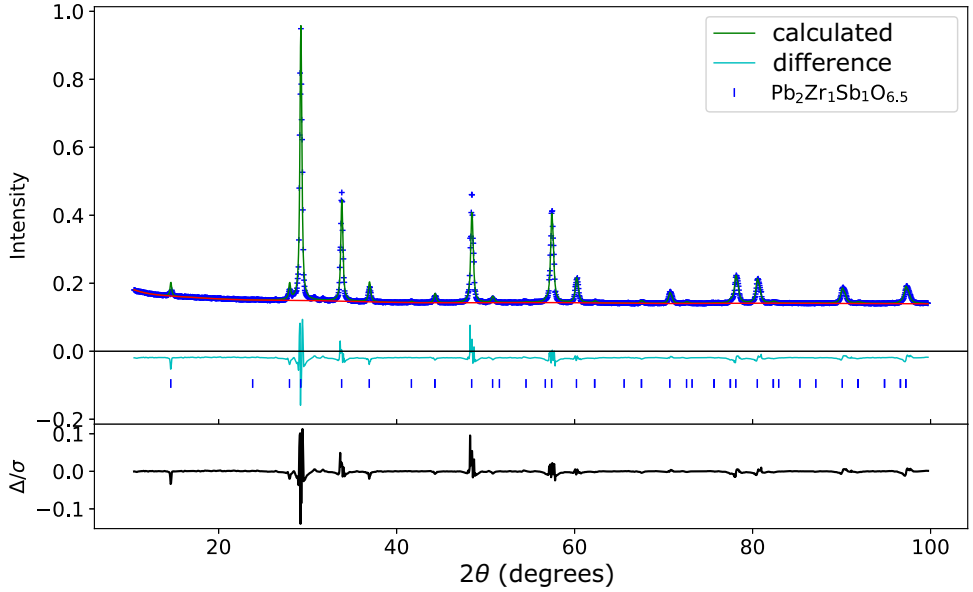

FIG. 15. Rietveld refinement of the pattern representing $Zr_2Sb_2Pb_4O_{13}$ against $Pb_2Zr_1Sb_1O_{6.5}$ (coll-62721), as reported in the ICSD.

[1] I. Levin, Inorganic Crystal Structure Database (ICSD), http://icsd.fiz-karlsruhe.de (2020).

[2] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, APL Mater. **1**, 011002 (2013).

[3] A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk, Scaling deep learning for materials discovery, Nature **624**, 80 (2023).

[4] N. J. Szymanski, B. Rendy, Y. Fei, R. E. Kumar, T. He, D. Milsted, M. J. McDermott, M. Gallant, E. D. Cubuk, A. Merchant, *et al.*, An autonomous laboratory for the accelerated synthesis of novel materials, Nature **624**, 86 (2023).

[5] J. Anwar, C. Leitold, and B. Peters, Solid-solid phase equilibria in the NaCl-KCl system, J. Chem. Phys. **152**, 144109 (2020).

[6] J. Huang, B. Ouyang, Y. Zhang, L. Yin, D.-H. Kwon, Z. Cai, Z. Lun, G. Zeng, M. Balasubramanian, and G. Ceder, Inhibiting collective cation migration in Li-rich cathode materials as a strategy to mitigate voltage hysteresis, Nat. Mater. **22**, 353 (2023).

[7] D. A. Keen and A. L. Goodwin, The crystallography of correlated disorder, Nature **521**, 303 (2015).

[8] I. Austin, C. Goodman, and A. Pengelly, New semiconductors with the chalcopyrite structure, J. Electrochem. Soc. **103**, 609 (1956).

[9] M. T. Anderson, K. B. Greenwood, G. A. Taylor, and K. R. Poeppelmeier, B-cation arrangements in double perovskites, Prog. Solid State Chem. **22**, 197 (1993).

[10] T. Graf, C. Felser, and S. S. Parkin, Simple rules for the understanding of Heusler compounds, Prog. Solid State Chem. **39**, 1 (2011).

[11] J. Klayman and Y.-W. Ha, Confirmation, disconfirmation, and information in hypothesis testing, Psychol. Rev. **94**, 211 (1987).

[12] B. Russell, *The Philosophy of Logical Atomism*, Routledge Classics (Routledge, Abingdon, 2009).

[13] C. Schlesinger, A. Fitterer, C. Buchsbaum, S. Habermehl, M. R. Chierotti, C. Nervi, and M. U. Schmidt, 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 **9**, 406 (2022).

[14] B. H. Toby and R. B. Von Dreele, GSAS-II: The genesis of a modern open-source all purpose crystallography software package, J. Appl. Crystallogr. **46**, 544 (2013).

[15] M. Serra, P. Salagre, Y. Cesteros, F. Medina, and J. E. Sueiras, Design of NiO-MgO materials with different properties, Phys. Chem. Chem. Phys. **6**, 858 (2004).

[16] S. Yagi, Y. Ichikawa, I. Yamada, T. Doi, T. Ichitsubo, and E. Matsubara, Synthesis of binary magnesium–transition metal oxides via inverse coprecipitation, Jpn. J. Appl. Phys. **52**, 025501 (2013).

[17] V. G. Tsirelson, A. S. Avilov, Y. A. Abramov, E. L. Belokoneva, R. Kitaneh, and D. Feil, X-ray and electron diffraction study of MgO, Acta Crystallogr. B **54**, 8 (1998).

[18] J. Singh, S. Lee, S. Kim, S. P. Singh, J. Kim, and A. K. Rai, Fabrication of 1D mesoporous NiO nano-rods as high capacity and long-life anode material for lithium ion batteries, J. Alloys Compd. **850**, 156755 (2021).

[19] H. Taguchi, S. Omori, M. Nagao, H. Kido, and M. Shimada, Crystal structure and magnetic properties of $(Ni_1 - xMg_x)_6MnO_8$, J. Solid State Chem. **118**, 112 (1995).

[20] Y. Guo, Y. Wei, H. Li, and T. Zhai, Layer structured materials for advanced energy storage and conversion, Small **13**, 1701649 (2017).

[21] C. Delmas, C. Fouassier, and P. Hagenmuller, Structural classification and properties of the layered oxides, Physica B+C **99**, 81 (1980).

[22] E. M. Seibel, J. Roudebush, H. Wu, Q. Huang, M. N. Ali, H. Ji, and R. Cava, Structure and magnetic properties of the α-$NaFeO_2$-type honeycomb compound $Na_3Ni_2BiO_6$, Inorg. Chem. **52**, 13605 (2013).

[23] S. H. Kim, S. J. Kim, and S. M. Oh, Preparation of layered MnO_2 via thermal decomposition of $KMnO_4$ and its electrochemical characterizations, Chem. Mater. **11**, 557 (1999).

[24] J. S. Gardner, M. J. Gingras, and J. E. Greedan, Magnetic pyrochlore oxides, Rev. Mod. Phys. **82**, 53 (2010).

[25] C. Cascales, I. Rasines, P. G. Casado, and J. Vega, The new pyrochlores $Pb_2(M_{0.5}Sb_{1.5})O_{6.5}$ (M = Al, Sc, Cr, Fe, Ga, Rh), Mater. Res. Bull. **20**, 1359 (1985).

[26] C. Cascales, J. Alonso, and I. Rasines, The new pyrochlores $Pb_2M(Sb)O_{6.5}$ (M = Ti, Zr, Sn, Hf), J. Mater. Sci. Lett. **5**, 675 (1986).

[27] A. Roy and B. H. Berrie, A new lead-based yellow in the seventeenth century, Stud. Conserv. **43**, 160 (1998).

[28] A. Marchetti, R. Saniz, D. Krishnan, L. Rabbachin, G. Nuyts, S. De Meyer, J. Verbeeck, K. Janssens, C. Pelosi, D. Lamoen, *et al.*, Unraveling the role of lattice substitutions on the stabilization of the intrinsically unstable $Pb_2Sb_2O_7$ pyrochlore: Explaining the lightfastness of lead pyroantimonate artists' pigments, Chem. Mater. **32**, 2863 (2020).

[29] S. Guillemet-Fritsch, J. Baudour, C. Chanel, F. Bouree, and A. Rousset, X-ray and neutron diffraction studies on nickel zinc manganite $Mn_{2.35-x}Ni_{0}.65Zn_xO_4$ powders, Solid State Ion. **132**, 63 (2000).

[30] L. Gama, C. O. Paiva-Santos, C. Vila, P. N. Lisboa-Filho, and E. Longo, Characterization of nickel doped $Zn_7Sb_2O_{12}$ spinel phase using Rietveld refinement, Powder Diffr. **18**, 219 (2003).

[31] A. Oleś, Neutron diffraction study of the $ZnFe_{2-x}Cr_xO_4$ series, Phys. Status Solidi (A) **3**, 569 (1970).

[32] T. Yamamoto, Y. Otsubo, T. Nagase, T. Kosuge, and M. Azuma, Synthesis and structure of vacancy-ordered perovskite $Ba_6Ta_2Na_2X_2O_{17}$ (X = P, V): Significance of structural model selection on discovered compounds, Inorg. Chem. (2024).

[33] M. C. Knapp and P. M. Woodward, A-site cation ordering in AA'BB'O_6 perovskites, J. Solid State Chem. **179**, 1076 (2006).

[34] K. Ji, K. N. Alharbi, E. Solana-Madruga, G. T. Moyo, C. Ritter, and J. P. Attfield, Double double to double perovskite transformations in quaternary manganese oxides, Angew. Chem. Int. Ed. **60**, 22248 (2021).

[35] Y. Wang, W. Li, G. Hu, Z. Peng, Y. Cao, H. Gao, K. Du, and J. B. Goodenough, Electrochemical performance of large-grained $NaCrO_2$ cathode materials for Na-ion batteries synthesized by decomposition of $Na_2Cr_2O_7{\cdot}2H_2O$, Chem. Mater. **31**, 5214 (2019).

[36] A. G. Nord, Neutron diffraction studies of $NiCoP_4O_{12}$ and $NiZnP_4O_{12}$, Mater. Res. Bull. **18**, 765 (1983).

[37] D. Wei, X. Yang, Y. Liu, and H. J. Seo, Structural and optical properties of $Mg_{1-x}Mn_xP_2O_6$ (x = 0–1.0) magnesium metaphosphate, J. Am. Ceram. Soc. **106**, 4246 (2023).

[38] A. Daidouh, C. Durio, C. Pico, M. Veiga, N. Chouaibi, and A. Ouassini, Structural and electrical study of the alluaudites $(Ag_{1-x}Na_x)_2FeMn_2(PO_4)_3$ (x = 0, 0.5 and 1), Solid State Sci. **4**, 541 (2002).

[39] S. Barth, R. Olazcuaga, P. Gravereau, G. Le Flem, and P. Hagenmuller, $Mg_0.5Ti_2(PO_4)_3$—a new member of the NASICON family with low thermal expansion, Mater. Lett. **16**, 96 (1993).

[40] M. Hidouri, N. Sendi, A. Wattiaux, and M. B. Amara, Structural study by X-ray diffraction and Mössbauer spectroscopy of a new synthetic iron phosphate: $K_4MgFe_3(PO_4)_5$, J. Phys. Chem. Solids **69**, 2555 (2008).

[41] S. T. Norberg, A. N. Sobolev, and V. A. Streltsov, Cation movement and phase transitions in KTP isostructures; X-ray study of sodium-doped KTP at 10.5 K, Acta Crystallogr. B **59**, 353 (2003).

[42] A. Boukhari, A. Moqine, and S. Flandrois, Synthesis and characterization of new copper (II) mixed diphosphates (M, $Cu)_2P_2O_7$ with M = Mg, Ca, Sr, and Ba, J. Solid State Chem. **87**, 251 (1990).

[43] D. Lenaz and F. Princivalle, Crystal chemistry of clinopyroxenes from Oligo-Miocene volcanics of Montresta (Sardinia, Italy), Mineral. Petrol. **90**, 157 (2007).

[44] S. D. Griesemer, L. Ward, and C. Wolverton, High-throughput crystal structure solution using prototypes, Phys. Rev. Mater. **5**, 105003 (2021).

[45] C. Li and J. Zhong, Highly efficient broadband near-infrared luminescence with zero-thermal-quenching in garnet $Y_3In_2Ga_3O_12{:}Cr^3+$ phosphors, Chem. Mater. **34**, 8418 (2022).

[46] S. N. Britvin, M. N. Murashko, M. G. Krzhizhanovskaya, N. S. Vlasenko, O. S. Vereshchagin, Y. Vapnik, and V. N. Bocharov, Crocobelonite, $CaFe^3_2+(PO_4)_2O$, a new oxyphosphate mineral, the product of pyrolytic oxidation of natural phosphides, Am. Mineral. **108**, 1973 (2023).

[47] M. Peplow, Robot chemist sparks row with claim it created new materials, Nat. News (2023).

[48] B. H. Toby, *R* factors in Rietveld analysis: How good is good enough?, Powder Diffr. **21**, 67 (2006).

[49] C. Meneghini, S. Ray, F. Liscio, F. Bardelli, S. Mobilio, and D. Sarma, Nature of "disorder" in the ordered double perovskite Sr_2FeMoO_6, Phys. Rev. Lett. **103**, 046403 (2009).

[50] R. A. Mayo, K. M. Marczenko, and E. R. Johnson, Quantitative matching of crystal structures to experimental powder diffractograms, Chem. Sci. **14**, 4777 (2023).

[51] A. Lunt, H. Fakhruldeen, G. Pizzuto, L. Longley, A. White, N. Rankin, R. Clowes, B. Alston, L. Gigli, G. M. Day, S. Y. Chong, and A. Cooper, Modular, multi-robot integration of laboratories: An autonomous workflow for solid-state chemistry, Chem. Sci. **15**, 2456 (2024).

[52] J. E. Salgado, S. Lerman, Z. Du, C. Xu, and N. Abdolrahim, Automated classification of big x-ray diffraction data using deep learning models, npj Comput. Mater. **9**, 214 (2023).

[53] M. Jaros, Electronic properties of semiconductor alloy systems, Rep. Prog. Phys. **48**, 1091 (1985).

[54] L. Bellaiche and D. Vanderbilt, Virtual crystal approximation revisited: Application to dielectric and piezoelectric properties of perovskites, Phys. Rev. B **61**, 7877 (2000).
[55] C. Chen, E. G. Wang, Y. M. Gu, D. M. Bylander, and L. Kleinman, Unexpected band-gap collapse in quaternary alloys at the group-III-nitride/GaAs interface: GaAlAsN, Phys. Rev. B **57**, 3753 (1998).
[56] N. J. Ramer and A. M. Rappe, Virtual-crystal approximation that works: Locating a compositional phase boundary in $Pb(Zr_{1-x}Ti_x)O_3$, Phys. Rev. B **62**, R743 (2000).
[57] R. Kikuchi, A theory of cooperative phenomena, Phys. Rev. **81**, 988 (1951).
[58] J. Sanchez, F. Ducastelle, and D. Gratias, Generalized cluster description of multicomponent systems, Phys. A: Stat. Mech. Appl. **128**, 334 (1984).
[59] A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, Special quasirandom structures, Phys. Rev. Lett. **65**, 353 (1990).
[60] R. Grau-Crespo, S. Hamad, C. R. A. Catlow, and N. H. de Leeuw, Symmetry-adapted configurational modelling of fractional site occupancy in solids, J. Phys.: Condens. Matter **19**, 256201 (2007).
[61] M. Yaghoobi and M. Alaei, Machine learning for compositional disorder: A comparison between different descriptors and machine learning frameworks, Comput. Mater. Sci. **207**, 111284 (2022).
[62] Z. Cai, *et al.*, Realizing continuous cation order-to-disorder tuning in a class of high-energy spinel-type Li-ion cathodes, Matter **4**, 3897 (2021).
[63] Z. Tan, Y. Peng, J. An, Q. Zhang, and J. Zhu, Critical role of order-disorder behavior in perovskite ferroelectric $KNbO_3$, Inorg. Chem. **60**, 7961 (2021).
[64] I. Mazin, Inverse Occam's razor, Nat. Phys. **18**, 367 (2022).
[65] H. T. Stokes, B. J. Campbell, and D. M. Hatch, FINDSYM, https://stokes.byu.edu/iso/findsym.php (2023).