



HAL
open science

Enumeration as a tool for structure solution – a materials genomic approach to solving the cation-ordered structure of $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$

Gerard S. Mattei, John M. Dagdelen, Matteo Bianchini, Alex M. Ganose, Anubhav Jain, Emmanuelle Suard, François Fauth, Christian Masquelier, Laurence Croguennec, Gerbrand Ceder, et al.

► To cite this version:

Gerard S. Mattei, John M. Dagdelen, Matteo Bianchini, Alex M. Ganose, Anubhav Jain, et al.. Enumeration as a tool for structure solution – a materials genomic approach to solving the cation-ordered structure of $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$. *Chemistry of Materials*, American Chemical Society, 2020, 32 (20), pp.8981-8992. 10.1021/acs.chemmater.0c03190 . hal-02959314

HAL Id: hal-02959314

<https://hal.archives-ouvertes.fr/hal-02959314>

Submitted on 12 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Enumeration as a tool for structure solution – a materials genomic approach to solving the cation-ordered structure of $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$

Gerard S. Mattei,^{a,b} John M. Dagdelen,^c Matteo Bianchini,^{c,†} Alex M. Ganose,^d Anubhav Jain,^d Emmanuelle Suard,^e Francois Fauth,^f Christian Masquelier,^{g,i} Laurence Croguennec,^{h,i} Gerbrand Ceder,^{c,d} Kristin A. Persson,^{c,d} Peter G. Khalifah^{a,b,*}

^a Department of Chemistry, Stony Brook University, Stony Brook, New York 11794-3400, USA

^b Chemistry Division, Brookhaven National Laboratory, Upton, New York 11793, USA

^c Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

^d Department of Materials Science and Engineering, University of California, Berkeley, Berkeley, California 94720, USA

^e Institut Laue-Langevin, 71 avenue des Martyrs, 38042 Grenoble Cedex 9, France

^f CELLS – ALBA synchrotron, Cerdanyola del Vallès, 08290 Barcelona, Spain

^g Laboratoire de Réactivité et de Chimie des Solides, Université de Picardie Jules Verne, CNRS-UMR#7314, F-80039 Amiens Cedex 1, France

^h CNRS, Univ. Bordeaux, Bordeaux INP, ICMCB UMR 5026, F-33600 Pessac, France

ⁱ RS2E, Réseau Français sur le Stockage Electrochimique de l'Energie, FR CNRS #3459, F-80039 Amiens Cedex 1, France

ABSTRACT: While powder diffraction methods are routinely utilized to optimize structural models for compounds whose crystal structures are known, the determination of unknown structures is far more challenging. When the unknown structure is large, structure solution can become a virtually intractable problem using standard structure solution methodologies, especially when the space group cannot be unambiguously resolved. One such system is the promising Na-ion battery cathode material $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ whose high temperature and room temperature structures were previously solved, but whose more complex low-temperature structure could not be determined. Here, a novel materials genomic approach is demonstrated for the solution of the unknown 100 K structure of $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ in which enumeration methods are first used to generate a large number (~3,000) of trial structures based on plausible orderings of Na ions and then automated Rietveld refinements are carried out to optimize each of these trial structures. Based on both the analysis of the ensemble of optimized trial structures and the density functional theory energy minimization of selected trial structures, the 100 K structure of $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ is best described as belonging to the space group $A2_1am$ with unit cell dimensions of $a = 9.01928(4)$, $b = 27.1379(1)$, $c = 10.73307(5)$. The 100 K unit cell has a large volume of $2627.07(2) \text{ \AA}^3$ with $Z = 12$ and 33 independent crystallographic sites (9 Na, 3 V, 3 P, 12 O, and 6 F) that is 3x and 6x larger than the room- and high-temperature polymorphs of this phase, respectively. The novel methods described here will be generally applicable for the solution of the complex cation-ordered structures that commonly occur for battery materials.

INTRODUCTION

The development of Li-ion batteries has had a transformative impact on our society, which was formally recognized with the award of the 2019 Nobel Prize in Chemistry to scientists who played a key role in unlocking this technology.¹ Li-ion batteries have been found to have the highest energy density among all viable rechargeable battery technologies developed to date,² resulting in the widespread adoption of hybrid electric vehicles and plug-in electric vehicles, as well as the stated goal of some countries to fully transitioning to electric vehicles over the next decade or so.³ However, there are some concerns about the availability of sufficient Li for the universal implementation of Li-ion batteries in the automotive sector, and as a result, there have been vigorous efforts to develop alternative Na-ion battery technologies in which the slightly lower energy densities are

counterbalanced by the greater abundance and lower cost of Na ions.⁴

One promising Na-ion cathode materials developed in recent years is $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ (NVPF), which has a large theoretical capacity (~256 mAh/g based on the exchange of four Na ions, although ~128 mAh/g (2 Na) are easily achievable and ~192 mAh/g (3 Na) can be achieved with wide voltage windows and custom cycling protocols), a high average discharge voltage and outstanding rate capability⁴⁻¹¹. Based on the analysis of variable-temperature synchrotron diffraction data, three structural polymorphs of NVPF have been identified. Above 395 K, the structure was determined to be tetragonal (space group $I4/mmm$, #139) with unit cell dimensions of $a = b = 6.39271(3) \text{ \AA}$ and $c = 10.75589(6) \text{ \AA}$ with a volume of $439.558(3) \text{ \AA}^3$. On cooling to room temperature, the symmetry is lowered to orthorhombic (space group $Amam$, #63) with lattice parameters of $a =$

9.02847(3) Å, $b = 9.04444(3)$ Å and $c = 10.74666(6)$ Å, as well as a doubled unit cell volume of $877.544(6)$ Å³.¹² The low-temperature structure of this phase (seen below 250 K) could not be determined in previous studies, though prior indexing studies indicated the likely presence of a supercell 3X larger in volume than the room temperature structure resulting from the tripling of the b -axis of that unit cell, as is summarized in Table 1.¹³

Table 1. Comparison of Na₃V₂(PO₄)₂F₃ polymorphs

T (K)	S.G.	a	b	c	Vol. (Å ³)
400 (HT)	<i>I4/mmm</i>	6.39	6.39	10.76	440 (1x)
300 (RT)	<i>Amam</i>	9.03	9.04	10.75	878 (2x)
100 (LT)	unknown	9.02	27.14	10.73	2627 (6x)

The room temperature (RT) structure of NVPF shares a common framework with the high-temperature (HT) structure of this phase. This framework is comprised of building blocks of VO₄F₂ octahedral dimers (with the F anions located at apical sites) and PO₄ tetrahedra, as illustrated in Figure 1. The primary difference between the HT tetragonal and the RT orthorhombic structure is the arrangement of Na ions, whose crystallographic sites are distributed about the VO₄F₂ octahedra with 4-fold rotational symmetry in the HT structure and with mirror symmetry in the RT structure when viewed along the (pseudo)tetragonal axis. Both the HT and RT structures contain partially occupied Na crystallographic sites, and the structure distortion associated with the HT->RT transition is driven by the partial localization of Na ions on cooling, leading to the breaking of the tetragonal symmetry.

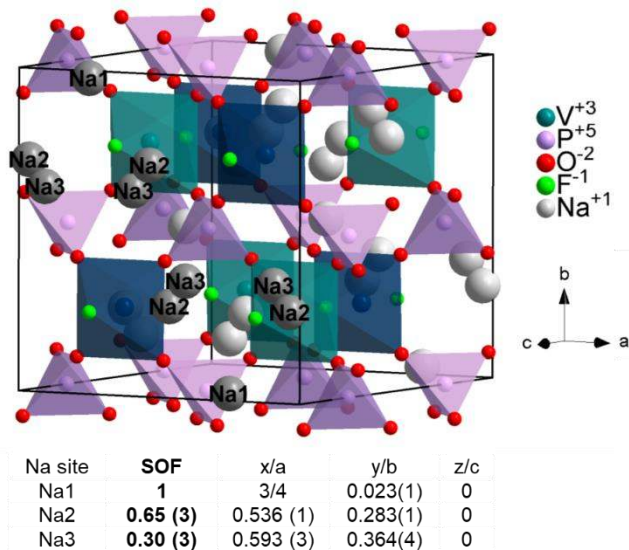


Figure 1. Known 300 K crystal structure of RT-NVPF, with symmetrically distinct sodium sites labeled. Coordination polyhedra are shown in lavender for P and in light and dark teal for V. Crystallographic site information for the three sodium sites (including their site occupancy factor, SOF) is also shown.

The close relationship between the RT and LT structures suggests that the RT->LT structure transition is also driven by the localization of Na ions on cooling. In the previously reported RT-NVPF structure, the Na1 site is fully occupied while the

Na2 and Na3 sites are about 2/3 and 1/3 occupied, respectively. The fact that the preliminary indexing studies show a tripling of the unit cell volume on cooling across the RT->LT transition strongly suggests that this transition is driven by the complete localization of Na ions that were distributed across Na2 and the Na3 sites in the RT structure. It is therefore expected that understanding the nature of the ordering of Na ions will be critical to solving the LT structure. This knowledge is required to understanding the crystal chemical interactions that drive the two-phase transitions and to understanding the preferred local coordination environments of Na ions that control the mobility of Na ions within this structure at all temperatures.

Despite the importance of knowing the LT-NVPF structure, its determination is an exceptionally difficult challenge that cannot readily be accomplished using conventional tools for the structural analysis of powder diffraction data. Density functional theory (DFT) studies have previously suggested that the ordering of Na ions can occur within the boundaries of the RT-NVPF unit cell.¹⁴ However, the DFT structure is not consistent with preliminary indexing studies carried out using high-resolution synchrotron powder diffraction data that strongly suggest that the LT phase has a C -centered orthorhombic cell with a volume 3x larger than the RT structure.¹³ Although the indexing result is robust, the small magnitude of the distortions causes the superstructure reflections to be ~100X weaker than the sub-cell reflections and for this reason does not allow clear discrimination between the 22 possible space group assignments consistent with the choice of a C -centered orthorhombic cell. Also, the 1:3 volume relationship between the LT and RT polymorphs suggests that the LT structure does not belong to a maximal subgroup of the RT structure and thus cannot be trivially deduced. All these challenges are further magnified by the large size of the LT cell (contains 216 atoms in a large volume of ~2600 Å³) and the fact that each atomic site will be nearly but not exactly symmetrically equivalent to other closely related sites.

For these reasons, we have chosen to take a novel materials genomic approach to structure solution in which a multitude of trial structures based on all plausible orderings of Na ions are explicitly enumerated and then individually tested against the experimental powder diffraction through automated Rietveld refinement. While enumeration methods are regularly combined with first principles DFT methods to identify the lowest energy structural polymorph associated with the ordering of chemical species in systems with many plausible configurations,¹⁴ to the best of our knowledge experimental powder diffraction data has never been utilized as the primary metric to assess the validity of large sets of trial configurations generated using modern enumeration methods¹⁵⁻¹⁶ though we do note one instance where powder diffraction data were used to pre-screen DFT energy minimization trials.¹⁷ Using this novel approach, a number of closely related structural models capable of reproducing the experimental observed powder diffraction patterns were obtained. This necessitated further efforts to identify the most correct structural model, and a combination of structural and density-functional theory analyses were used to suggest an optimal model for the low-temperature structure of Na₃V₂(PO₄)₂F₃, providing the first structural model capable of completely describing its experimentally observed supercell.

EXPERIMENTAL SECTION

$\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ was synthesized as described in previous publications.¹² Samples were enclosed in borosilicate capillaries (0.7 – 1.0 mm in diameter). Synchrotron X-ray powder diffraction patterns were measured at the MSPD beamline of the ALBA Synchrotron (Barcelona)¹⁸ utilizing the high-angular mode of the station (13 channel Si 111 multi-analyzer setup) where data are collected in the Debye-Scherrer geometry over a 2θ angular range of $2^\circ - 50^\circ$ at a wavelength of 0.62003 Å. Temperature control was provided by a Cryostream Plus nitrogen blower from Oxford Cryosystems. Both the Jana2006 and TOPAS software packages were used to identify symmetry and systematic absences in the 100 K XRD pattern of LT-NVPF.

Trial structural generation and analysis was carried out using pymatgen, the open source Python library for implementing and manipulating crystal structures developed by the Materials Project.¹⁹ Structure enumeration was performed using a pymatgen interface to the external enumlib library,^{16, 20} with a different atom type used as a placeholder to denote each choice for a local environment for later substitution. Pawley and Rietveld fits of all enumerated trial structures for the 100 K data were performed using a custom-built Python interface to the command line implementation of the TOPAS software (v6).²¹ Additional custom Python scripts were used to automatically generate TOPAS input files for each trial structure and to extract refined parameter values from output files.

Rietveld refinements of trial structures investigating their ability to reproduce the 100 K diffraction data were performed using two refinement strategies, one static and one dynamic. In the static one, the atomic coordinates returned by the enumeration were not allowed to vary from their initial position, while in the dynamic one, all atomic positions were freely refined within the crystallographic constraints of the space group symmetry of each trial structure resulting from enumeration. In all refinements, isotropic atomic displacement parameters (ADPs) were allowed to vary but were constrained to be equal for all sites containing a given elemental type.

In order to maximize the sensitivity to structural features, care was taken to optimize the fitting of the experimental diffraction data. Lattice parameters were refined along with the inclusion of a 2θ zero error correction. Peak shapes were modeled with pseudo-Voigt shapes with an additional exponential convolution to fit their asymmetry. Different breadths were used to model subcell ($k = 3n$ when the b -axis was tripled with regard to that of the RT structure) and supercell reflections following a complex aphysical model (sample input file included as Supporting Information) whose goal was to optimize peak shape predictions. It is expected that the HKL-dependent broadening was due to different coherence lengths of subcell and supercell peaks while the asymmetry was due to strain gradients.

Density functional theory (DFT) as implemented in the Vienna Ab Initio Simulation Package (VASP) was used to carry out structure relaxation.²²⁻²³ Exchange-correlation effects on the electron density were accounted for using the Perdew-Burke-Ernzerhof functional.²⁴ The recommended “Na_pv”, “V_pv”, “P”, “O”, and “F” pseudopotentials distributed in VASP were employed for all calculations. In addition, corrections for the Hubbard U were implemented for vanadium ($U = 3.25$ eV) using the standard Materials Project values as documented and implemented in the pymatgen VASP input sets.¹⁹ An energy cutoff of 520 eV and an automatically generated, Γ -centered k -

point mesh with a reciprocal density of 64 k -points per Å⁻³ were used, as implemented in the base structure optimization workflow defined in the atomate package.²⁵ Symmetry constraints were explicitly turned off during the structure relaxation process. Atomic displacement parameters were computed using the phonopy package²⁶ following the methodology described by Grosse-Kunstleve and Adams.²⁷ Second order interatomic force constants were calculated from finite displacements in a $1 \times 1 \times 2$ supercell of the 108 atom primitive cell. Forces were obtained using the PBE exchange correlation functional using a tight energy convergence criterion of 10^{-6} eV and a single k -point at Γ .

RESULTS

Of the three known polymorphs of $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$, an important battery cathode previously investigated through synchrotron powder diffraction studies, only the structure of the low-temperature polymorph (LT-NVPF) remains unknown. Unlike the RT and HT polymorphs, the LT structure is expected to have fully ordered (and thus fully occupied) Na sites and will provide clear insights into the local coordination environments that are largely responsible for determining the activation energies for Na ion motion. Although the quality of powder diffraction data for the LT phase was comparable to that of the RT and HT polymorphs, previous attempts at structure solution were unsuccessful primarily due to the difficulty of determining the space group symmetry of this phase, a consequence of the large volume of its unit cell and the very weak intensities of the superstructure peaks associated with the presumed Na ion ordering (Figure 2). Other than the expected contraction due to cooling, the LT diffraction pattern is virtually indistinguishable from the RT pattern until viewed at a high magnification to resolve the very weak superstructure peaks.

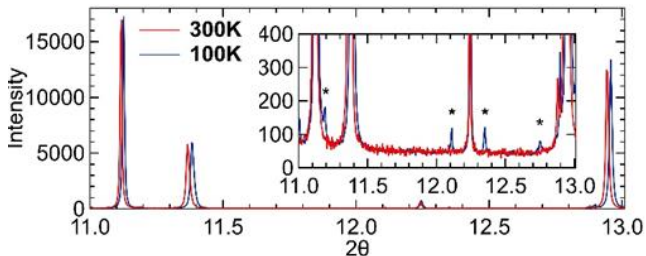


Figure 2. Comparison of synchrotron XRD data ($\lambda = 0.62$ Å) collected on NVPF at 300 K (RT, red) and 100 K (LT, blue). Inset is shown at an approximately 50X zoom, highlighting additional low-intensity supercell peaks associated with Na cation ordering (*).

Enumeration Strategy. Since conventional methods of structure solution did not give the structure of LT-NVPF, we chose to pursue a novel route for structure solution based on a materials genomics approach. For this, we explicitly enumerated and tested all plausible orderings of Na ions in the RT-NVPF compatible with two constraints. The first is a unit cell where the b -axis of the RT cell is tripled (Table 1), as suggested by the preliminary indexing of the LT-NVPF diffraction data of Figure 2. Although a centered lattice was suggested by preliminary indexing, we did not enforce this constraint due to the inability to make definitive conclusions based on the weak superstructure peaks. The second is the presence of three different ordering motifs, as suggested by the 2:1 ratio of disordered Na site occupancies in the RT structure.

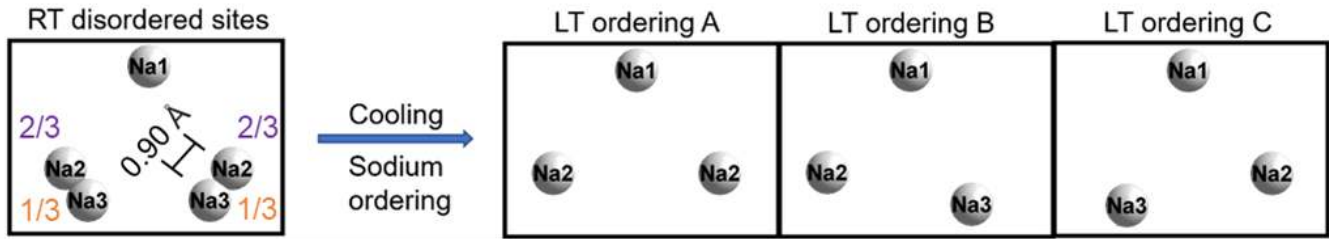


Figure 3. Clustering of partially occupied Na2 and Na3 sites in RT-NVPF structure at 300 K (left) and the corresponding three equally likely ordered arrangements (A, B, C) of fully occupied Na2 and Na3 sites hypothesized to occur at 100 K for LT-NVPF (right).

There are three types of Na crystallographic sites in the room temperature structure of NVPF, with a cluster of neighboring Na positions shown on the left side of Figure 3. The Na1 (4c) atomic site is fully occupied and is well-separated from the other Na sites and was therefore judged to be essentially unaffected by the RT->LT structural transition. In contrast, there are two pairs of Na2 and Na3 sites in the cluster. The Na2 (8f) and Na3 (8f) sites are partially occupied with experimentally determined fractional occupancies that are about 2/3 and 1/3, respectively, and have a summed occupancy in the cluster that is double that of the Na1 site when accounting for their higher multiplicity. Since the Na2 and Na3 sites are separated from each other by only 0.90 Å (a distance much less than the diameter of Na ions) it is impossible for both sites in a pair to be simultaneously occupied. It is therefore expected that the RT->LT structural transition is driven by Na cations ordering in a pattern that results in each Na2- and Na3-derived site being either fully occupied or completely empty.

Both the preliminary indexing results for the LT structure (in which the cell volume triples) and the refined site occupancies (which can be described as fractions with a denominator of 3) suggest that there are three different types of structural motifs that can result from the ordering of Na atoms. Based on the ratio of site occupancies between the Na2 and Na3 site (2:1 ratio), it was hypothesized that the three motifs occur in equal proportions and have occupancy patterns of (A) Na2 / Na2, (B) Na2 / Na3, or (C) Na3 / Na2, where the first and second label indicate the atomic site that is occupied on the left and right side of the cluster, respectively, as illustrated in Figure 3. An occupancy pattern of Na3 / Na3 is chemically implausible due to the very short distance (2.8 Å) between Na ions in this pair.

Even under the constraining assumption that the LT-NVPF structure can be described as an ordering pattern of these three structural motifs, there are an immense number of possible configurations due to the large unit cell size. Within the tripled cell volume, there are $3^{12} = 531,441$ possible permutations of sodium orderings. However, the use of symmetry-aware software which can identify and eliminate duplicate structures which are symmetry equivalents^{16, 20} allowed the identification of a more tractable subset of 2,940 unique orderings. The significant challenge of investigating all of these trial structures was made possible through the development of automated scripts for high-throughput Rietveld refinement and the use of the powerful TOPAS software package for Rietveld refinement which enabled the rapid convergence of refinements for each trial structure and allowed all refinements to be carried out in ~55 h on a laptop personal computer.

Rietveld Evaluation of Trial Structures. While enumeration methods have commonly been applied to generate trial structures for problems associated with density functional the-

ory (DFT) such as the use of cluster expansion methods to understand the local configurations within structures with site disorder,²⁸⁻³² they have not to the best of our knowledge been previously applied to purely experimental structure solution and refinement problems. There is a close analogy between DFT and structure solution problems in that they both are designed to minimize a single parameter – energy in the case of DFT and R_{wp} in the case of structure refinements. However, diffraction experiments have the additional challenge of working with data with experimental uncertainties – the structure with the lowest R_{wp} need not be the most correct structure due to effects of these errors as well as to the different number of refinement parameters associated with different trial structures.

A schematic illustration of the present structure solution problem is given in Figure 4. The enumeration process will provide many trial structures which reside on a complex landscape of R_{wp} values obtained with respect to the experimental data. Each trial structure can then be improved through the process of Rietveld refinement in which the coordinates of atomic sites are optimized, driving the structure to the vicinity of either a local or a global minimum in R_{wp} after the optimization process converges. The number of potentially accessible local minima and the ability to find the global minimum will depend on the complexity of the system, the capabilities of the optimization algorithm, and the diversity and reasonableness of the starting trial structures.

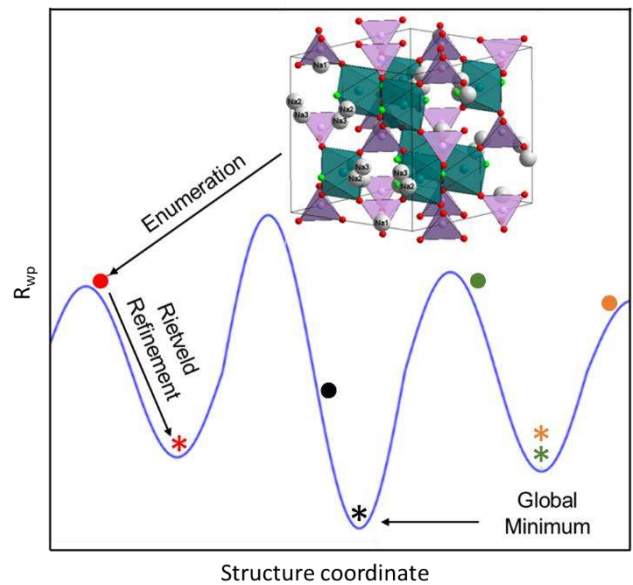


Figure 4. Schematic illustrating the structure solution process. All plausible Na ordering patterns (2,940) were generated through an enumeration process. Each trial structure (circle) was then opti-

mized through Rietveld refinement to minimize R_{wp} . After this refinement (star, shown with same color), the optimized trial structures should exist as either local or global minima.

In order to get a sense for the energy landscape associated with the LT-NVPF structure, structure refinements were carried out in two ways. First, refinements were carried out with atomic coordinates fixed to the initial values returned through the enumeration process (Figure 5a). Second, refinements were carried out with the atomic coordinates of all crystallographic sites freely refined within symmetry constraints (Figure 5b) associated with the initial trial structure. For these plots, the small subset of enumerated structures having an A -centered lattice (18 structures, further details in Figure S1) is indicated with red symbols. A summary of the enumeration and refinement results organized by space group is provided in Table 2.

With fixed atomic coordinates, the fits using the as-returned enumerated structures vary nearly continuously across a wide range of R_{wp} values ranging from about 11 to 15% with three

very broad peaks in this distribution (Figure S2). Enumerated structures with an A -centered lattice were responsible for some of the best fits, supporting the validity of this assignment from the preliminary indexing studies. However, there was no single best structure that was clearly superior to all others.

Upon refinement of atomic coordinates, the distribution of R_{wp} values across the 2,940 trial structures resulted in most of the structures being organized into about half a dozen narrow bands that presumably reflect key minima, as previously suggested in Figure 4. Surprisingly, most (~2,500) trial structures after optimization were found to reside in a sharp band at a R_{wp} value of about 9.5% that appears to correspond to the global minimum. This indicates that the biggest challenge is not finding the one structural model that effectively fits the diffraction data but instead discriminating between many refined structures that can effectively fit the data in order to determine which one space group symmetry best describes the LT-NVPF structure.

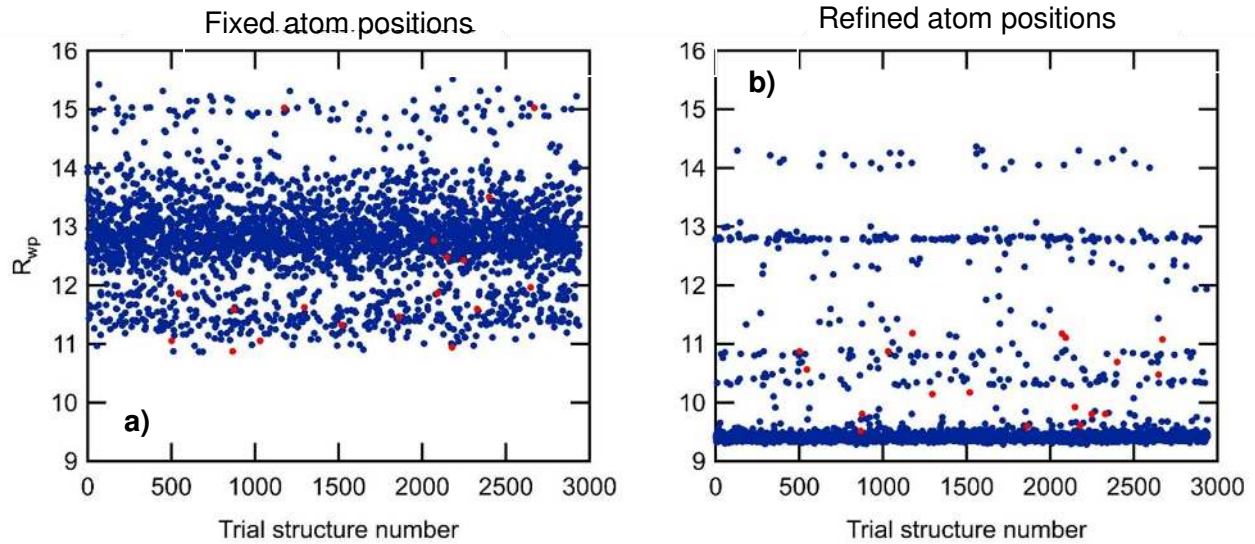


Figure 5. Comparison of R_{wp} values for all 2,940 trial structures obtained after Rietveld refinements in which atom positions were fixed (left) or were freely refined, allowing structures to be optimized (right). Refinements in centered space groups are indicated with red symbols, and are shown in more detail in Figure S1, while R_{wp} histograms are shown in Figure S2.

Table 2. Enumeration trial structure symmetry and refinement summaries.

S.G. #	Space group (non-standard)	Centering	a -axis sym.	b -axis sym.	c -axis sym.	-1	S.G. (std.)	# of trial models	# xyz params.	R_{wp} (mean)	R_{wp} (best)
36	$A 2_1 a m$	A	2_1	a, n	$2_1/m, b$		$Cmc2_1$	6	87	9.81	9.52
12	$A 1 1 2/m$	A	1	1	$2/m, b$	+	$C2/m$	6	86	10.97	10.57
58	$P2_1/n 2_1/n 2_1/m$	P	$2_1/n$	$2_1/n$	$2/m$	+	$Pnmm$	6	86	10.33	10.31
62	$P2_1/n 2_1/a 2_1/m$	P	$2_1/n$	$2_1/a$	$2_1/m$	+	$Pnma$	6	87	12.55	12.40
31	$P 2_1 n m$	P	2_1	n	m		$Pmn2_1$	68	174	9.64	9.29
26	$P 2_1 a m$	P	2_1	a	m		$Pmc2_1$	54	174	9.42	9.30
10	$P 1 1 2/m$	P	1	1	$2/m$	+	$P2/m$	36	172	10.14	9.46
11	$P 1 1 2_1/m$	P	1	1	$2_1/m$	+	$P2_1/m$	36	174	10.16	9.50
8	$A 1 1 m$	A	1	1	m, b		Cm	6	174	10.30	9.81
6	$P 1 1 m$	P	1	1	m		Pm	2716	348	9.72	9.33

Identification of Optimal Symmetry. The 2,940 trial structures identified through enumeration all belong to one of ten possible space groups, as listed in Table 2. We have chosen to use non-standard space group settings to describe these com-

pounds in order to permit a common orientation of cell axis relative to the RT structure (which itself has been described in the non-standard setting $Amam$ of space group #63 to preserve the c -axis alignment with the HT tetragonal structure). This enables

the relationship between common symmetry elements to be better illustrated. These ten space groups are either monoclinic with the c -axis unique or orthorhombic, and in either case can be primitive or A -centered. The enumerated structures are predominantly monoclinic (2,800), with the vast majority of these (2,716) belonging to the minimal symmetry space group $P11m$ which has only mirror planes in the Na ion layers. A total of 140 orthorhombic structures were identified. Only 18 structures have the centered lattice suggested by indexing studies with 6 of these in the monoclinic space group $A11m$, 6 in the monoclinic space group $A112/m$, and 6 in the orthorhombic space group $A2_1am$ (remembering that the C -centered lattice from indexing can be trivially transformed to an A -centered lattice).

The ten space groups in Table 1 are organized in groups based on the number of refined parameters for their crystallographic sites (fractional coordinates x , y , and z) that were optimized during refinement. The highest symmetry structures have 86 or 87 parameters, the second tier of compounds has double that, and the last tier ($P11m$) has quadruple that. While the best fits to the experimental data (as judged by R_{wp} values) were obtained using the lowest symmetry structures, this does not directly indicate the correctness of structures since those refinements had more free parameters. The most promising space group instead is $A2_1am$, which achieves excellent fits to the data with a minimal number of parameters. The three best fitting $A2_1am$ trial structures (#867, $R_{wp} = 9.52$; #1860, $R_{wp} = 9.60$; #2179, $R_{wp} = 9.61$) all fall in the lowest bands of Figure 5b and Figure S2 and are near the minimum R_{wp} value achieved across all trial structures (#840, $R_{wp} = 9.29$) which belonged to space group $P2_1nm$.

Further assessments of the relative merits of the ten space group choices can be made by comparing the symmetry elements present within each space group (Table 2) and their effect on fit quality. The five space groups containing inversion symmetry (-1) give the worst fits to experimental data as judged by their mean R_{wp} values, strongly indicating the absence of inversion symmetry in the true structure. An examination of the remaining symmetries indicates that the best fits ($R_{wp} \leq 9.33$) all contain a subset of the symmetry elements present in space group $A2_1am$ and further support the choice of this space group.

Curiously, trial structures with $A11m$ symmetry gave worse fits to the data than the optimal $A2_1am$ structure despite $A11m$ being a subgroup of $A2_1am$ and thus having only a subset of the $A2_1am$ symmetry constraints. We attribute this to these refinements universally getting stuck in local minima. This likely occurs because the enumerated Am structures necessarily contain features which violate $A2_1am$ symmetry, and because the six trial structures had insufficient structural diversity to relax to the global minimum during refinements. This hypothesis was explicitly tested by taking the optimal enumerated $A2_1am$ structure and then relaxing the symmetry down to $A11m$ prior to the start of Rietveld refinement. When this was done, the refinement in $A11m$ gave a low R_{wp} of 9.54 whose fit quality is indistinguishable from that obtained in $A2_1am$.

Further assessments of the relative merits of different space group symmetries can be obtained by visually examining the quality of the fit obtained using the best trial structure within each space group, as is shown on a log scale in Figure 6. The trial structures for space groups $Pnam$ and $Pnmm$ miss one or two of the largest supercell peaks and are clearly incorrect. The best trial structure for the presumed space group of $A2_1am$ does not miss any of the three clearly resolved supercell peaks in this region, and additionally, does not predict intensity where no

peaks are present in the data. The best fitting trial structures from the remaining space groups all predict some intensity in regions where there are no clear peaks in the experimental diffraction data as well as predicting a large number of peaks with negligible intensity. As such, the data support the space group of $A2_1am$ as the optimal choice.

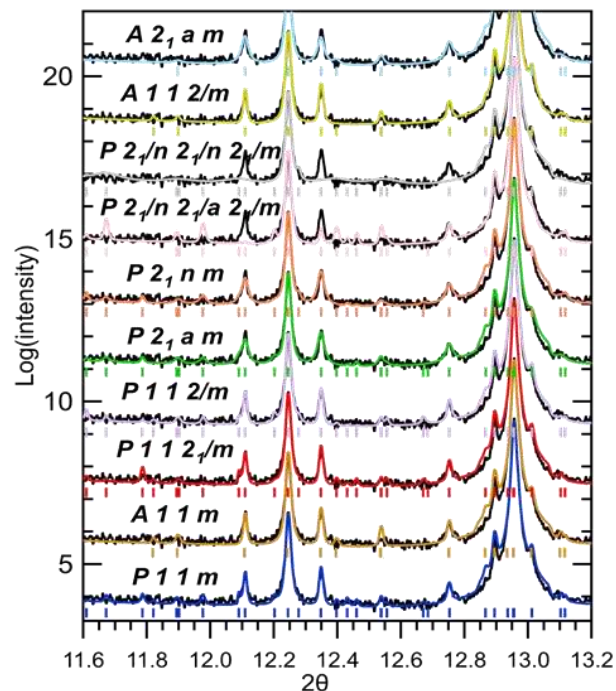


Figure 6. Comparison of best fits obtained for structures optimized by Rietveld refinement for each of the 10 space groups identified for enumerated structures in a region that allows superstructure peak intensities to be evaluated. Experimental data are shown in black and simulated intensities are shown as colored lines, with a vertical log scale that is offset by 1.85 for successive data sets.

A last experimental check of the suitability of the $A2_1am$ space group is the direct comparison of the best fit structure in this higher symmetry space group to the global best fit across all structures, which occurred in the lower symmetry space group $P2_1nm$. This comparison is shown in Figure 7 on a log scale in a region where the superstructure peaks can be most readily discerned. The space group $P2_1nm$ provides no discernable advantages in predicting experimentally observed peaks which are forbidden in $A2_1am$, nor does this space group have a smaller number of predicted peaks which are experimentally absent. The main advantage of $P2_1nm$ over $A2_1am$ appears to be in a slightly improved quality of fit to the experimental peak intensities – behavior that can be attributed to overparameterization of the refinement. An examination across the entire range of the fit of the LT-NVPF diffraction data to the refined $A2_1am$ structural model (Figure S3, with data provided as Supporting online information) shows that the intensities of the hundreds of clearly resolvable diffraction peaks in this pattern can be fit with only minor intensity errors. As such, the optimal $A2_1am$ structural model must closely match the true NVPF structure.

As a complement to the experimental tests of symmetry, the suitability of the $A2_1am$ space group symmetry was also assessed through DFT methods which were utilized to relax the best experimentally determined $A2_1am$ structure without applying any symmetry constraints ($P1$ symmetry). If the proposed

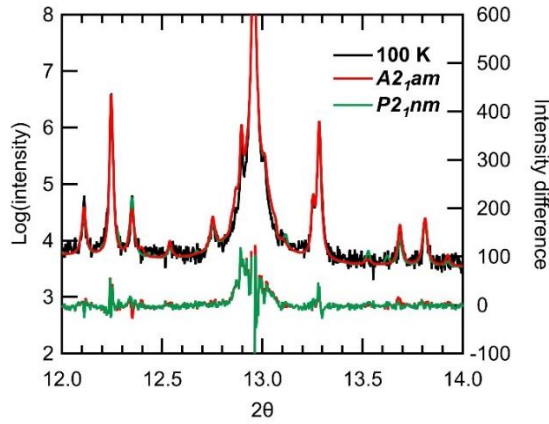


Figure 7. Comparison of the best fit in centered orthorhombic space group $A2_{1am}$ ($R_{wp} = 9.52$) with the best overall fit among all trials (space group $P2_{1nm}$, $R_{wp} = 9.29$). Data (black) and models are shown on a log scale; difference patterns (below) on a linear scale.

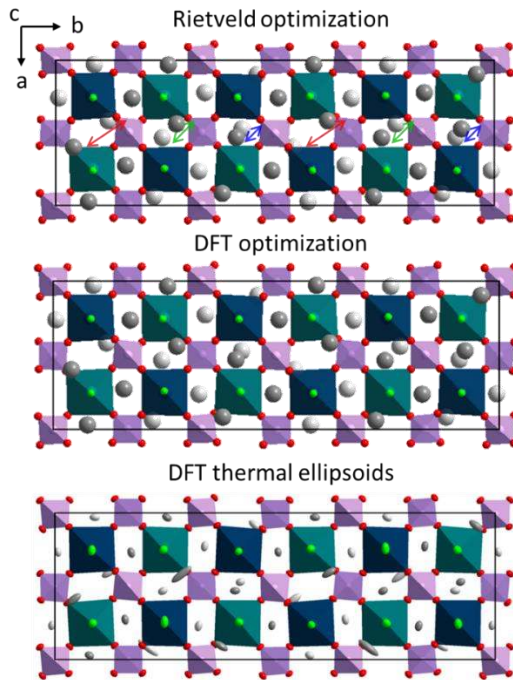


Figure 8. Polyhedral representation of optimal $A2_{1am}$ structure of LT-NVPF obtained after Rietveld refinement (top) and after DFT structure relaxation (center). The large (red), medium (blue), and small (green) gaps between Na ions in the center row are indicated by arrows. An additional plot of the second DFT calculation used to obtain atomic displacement parameters is also shown (bottom) with the thermal ellipsoids drawn at the 90% level. The VO_6 octahedra are shown in dark blue when the apical F is in the Na plane and in teal otherwise, and the PO_4 tetrahedra are shown in mauve. Na atoms are shown in dark grey in the upper layer and light grey in the lower layer, O are shown in red, and F in green.

$A2_{1am}$ space group symmetry is too high, the DFT energy optimization should result in relaxation to a lower symmetry space group. This did not occur during testing, and the DFT-relaxed structure was still best described by the $A2_{1am}$ space group. As shown in Figure 8, a direct comparison of the experimentally refined structure (top) and the DFT-relaxed structure

(center) shows no significant differences, again supporting the assignment of $A2_{1am}$ symmetry for LT-NVPF.

One curious aspect of the RT- \rightarrow LT structural transition of NVPF is that the unit cell volume triples as a result of this transition despite the absence of order 3 symmetry elements in the parent RT structure ($Amam$ symmetry). An examination of the optimal $A2_{1am}$ structure model in Figures 8 and S4 shows how this tripling develops. The Na atoms corresponding to the Na2 and Na3 sites in the RT structure align between vanadium octahedra in pairs with large (L), medium (M), or small (S) separations between the ions, and repeat in the pattern LMS moving along the central row of this structure as indicated by the red, blue, and green arrows, respectively, of Figure 8. It should be noted that this same type of sequence is generally observed for optimized trial structures that fall in the band with the lowest R_{wp} (Figure S5), though in a manner that is potentially changed by an origin shift, by axis permutations, or by the application of symmetry elements present in the structure, making it challenging to directly compare different trial structures. The robust observation of local bonding changes that drive the tripling of the b -axis in the original $Amam$ cell even in lower symmetry trial refinements strongly supports the assignment of $A2_{1am}$ symmetry.

Analysis of $A2_{1am}$ LT-NVPF structure. Details of the data collection are provided in Table 3, while crystallographic site information is provided in Table 4. Additionally, CIF files are provided as Supporting Information for the experimental structure, the DFT-relaxed structure, and a rescaled version of the DFT-relaxed structure which was isotropically expanded to have the same volume as the experimental cell. A comparison of key bond distances in these three structures is provided in Table 5. The experimentally refined LT-NVPF structure is

Table 3. Data collection parameters for LT-NVPF (100K)

Empirical formula	$Na_3V_2(PO_4)_2F_3$
Z	12
Molar mass (g/mol)	5013.48
Theoretical density (g/cm ³)	3.17
Space group	$A2_{1am}$ (#36)
Crystal system	Orthorhombic
Lattice parameters (Å)	$a = 9.01928$ (4) $b = 27.13787$ (12) $c = 10.73307$ (5)
Unit cell volume (Å ³)	2627.07 (2)
Radiation type	X-ray
Wavelength (Å)	0.62003
R_{wp}	9.52
d_{max}, d_{min}	$8.87 > d > 0.6155$
Unique reflns ($I/\sigma \geq 8$; all)	718; 3235
Temperature (K)	100

Table 4. LT-NVPF (100 K) crystallographic sites

Site	Wyck.	x	y	z
Na1	4a	0	0.0409(3)	0.4001(10)

Na2	4a	0	0.0751(4)	0.0263(14)
Na3	4a	0	0.1598(5)	0.2751(13)
Na4	4a	0	0.1740(5)	0.7363(13)
Na5	4a	0	0.2757(3)	0.5604(10)
Na6	4a	0	0.2568(4)	0.9654(11)
Na7	4a	0	0.3800(3)	0.0823(10)
Na8	4a	0	0.4120(4)	0.4678(12)
Na9	4a	0	0.4927(6)	0.2397(13)
V1	8b	0.1826(6)	0.41566(17)	0.7527(4)
V2	8b	0.1856(5)	0.74860(19)	0.7521(4)
V3	8b	0.1840(6)	0.08176(19)	0.7488(4)
P4	8b	0.2446(8)	0.5002(2)	-0.0041(7)
P5	8b	0.2439(6)	0.1681(3)	0.0016(10)
P6	8b	0.2524(5)	0.1669(3)	0.5019(10)
O1	8b	0.1539(14)	0.8019(5)	0.6011(16)
O2	8b	0.1642(17)	0.4697(5)	0.9004(18)
O3	8b	0.1651(17)	0.4667(5)	0.5903(19)
O4	8b	0.1609(16)	0.1321(5)	0.9134(16)
O5	8b	0.1697(15)	0.1387(5)	0.6073(16)
O6	8b	0.1609(16)	0.1984(5)	0.3970(17)
O7	8b	0.1673(18)	0.3666(5)	0.9052(17)
O8	8b	0.1708(16)	0.0355(5)	0.5841(16)
O9	8b	0.1540(14)	0.3018(5)	0.0889(16)
O10	8b	0.1626(14)	0.3631(5)	0.6036(15)
O11	8b	0.1714(16)	0.3033(5)	0.4025(17)
O12	8b	0.1655(14)	0.0365(4)	0.9107(14)
F1	8b	0.1329(17)	0.4204(5)	0.2490(11)
F2	8b	0.1243(11)	0.0852(5)	0.2395(9)
F3	8b	0.1376(14)	0.2458(5)	0.7419(10)
F4	4a	0	0.0856(7)	0.7394(13)
F5	4a	0	0.2481(8)	0.2543(19)
F6	4a	0	0.4146(6)	0.7404(14)

All sites were refined with B_{iso} values (in \AA^2) constrained to be equal for all sites for a given element, with refined values of Na: 0.60(3), V: 0.068(1), P: 0.093(2), O: 0.00(3), and F: 0.15(4). These values are uniformly low due to the absence of a correction for absorption and are generally lower by ~ 0.65 than the average B_{eq} values obtained from DFT calculations: Na: 1.33, V: 0.50, P: 0.52, O: 0.85, F: 0.88. The full DFT results are provided as Supporting Information in the form of a CIF file.

remarkable in its complexity in that it contains 33 unique crystallographic sites (9 Na, 3 V, 3 P, 12 O, and 6 F) whose fractional coordinates xyz require 87 parameters to describe. This structure is far more complex than that of the room-temperature polymorph which had only 9 crystallographic sites (3 Na, 1 V, 1 P, 2 O, 2 F) and 16 refined parameters associated with the fractional coordinates of these sites.

The large number of crystallographic sites in the LT-NVPPF structure combined with the correlations associated with the very similar coordination environment of independent crystallographic sites (that are derived from the same site in the higher-

symmetry RT / HT structures) and the difficulty of accurately quantifying the intensity of the weak superstructure peaks severely limits the accuracy with which bond distances can be determined. The DFT-optimized structure of this phase will not suffer from these experimental limitations and therefore provides an important reference that can give insights into both the true local geometries and the magnitude of the experimental errors. However, comparisons between DFT and experimental results can only be effectively made after adjusting the volume of the DFT-optimized structure to match the experimentally measured volume of the LT-NVPPF structure due to the well-known deviations between these two methods.^{33,34} Once this is done, it can be seen that the spans of P-O distances (1.54 – 1.56 \AA) and V-O distances (1.99 – 2.07 \AA) observed for the adjusted DFT data are much narrower than for the experimental structure obtained from Rietveld refinement against powder diffraction data (1.47 – 1.60 \AA and 1.90 – 2.08 \AA). These results are consistent with an uncertainty of about 0.05 \AA in the experimentally determined bond distances, which is about an order of magnitude larger than is typically seen when refinements are carried for modest size unit cells (such as that of RT-NVPPF) using similar quality powder diffraction data.

From thermodynamic arguments, it can be expected that the bonding preferences of the ions are more effectively satisfied in the LT polymorph than in the RT polymorph, especially with regard to the Na ions that are involved in an order-disorder transition. Within the RT polymorph, the Na1 site adopts a mono-capped trigonal prismatic environment (CN = 7) with a formula of NaO_4F_3 , with the four oxygens making up the corners of the prismatic face that is capped. The Na2 site shares this NaO_4F_3 mono-capped trigonal prismatic coordination environment (CN = 7), while the Na3 coordination differs from the other two sites only in missing the capping F ion of the trigonal prism (NaO_4F_2 , CN = 6). All three types of trigonal prisms in the RT-NVPPF structure are constructed from three pairs of atoms belonging to VO_4F_2 octahedra with half the atoms coming from octahedra located in the plane above the Na site (+z) and the other half located in the plane below the Na site (-z). In the case of the Na1 and Na2 sites, the two coordinating oxygens at the same height along z come from the edge of a single VO_4F_2 octahedron while in the case of the Na3 site, the two coordinating oxygens come from the corners of two different octahedra. The pair of F atoms on the edge of the trigonal prisms always come from the terminal apical positions of two separate VO_4F_2 octahedra (one above and one below), while the capping F atom is same plane as the Na sites and comes from the central apical position of a corner-sharing pair of VO_4F_2 octahedra. Due to the displacement of the Na3 site away from the center of the VO_4F_2 octahedral edge, this site is further away from the potential capping ligand (Na-F distance of ~ 3.2 \AA), resulting in the lower CN for this site.

Table 5. Bond distances in LT-NVPF from XRD, DFT, and by DFT linearly rescaled to the XRD cell volume

Site		XRD	DFT	DFT (adj.)	Site ¹	Ligand	XRD	DFT	DFT (adj.)	BVS ² (bond)
P1	1 x O	1.47(2)	1.545	1.521	Na1	2 x F	2.31(1)	2.365	2.330	0.18
	1 x O	1.50(2)	1.548	1.525	“3”	2 x O	2.48(2)	2.531	2.493	0.16
	1 x O	1.55(2)	1.548	1.525		2 x O	2.75(2)	2.696	2.655	0.08
	1 x O	1.58(2)	1.558	1.535	Na2	2 x O	2.31(2)	2.454	2.417	0.25
P2	1 x O	1.49(2)	1.551	1.528	“2”	2 x F	2.36(1)	2.434	2.398	0.16
	1 x O	1.54(2)	1.547	1.524		2 x O	2.53(2)	2.584	2.545	0.14
	1 x O	1.55(2)	1.551	1.528		1 x F	2.60(2)	2.616	2.577	0.08
	1 x O	1.60(2)	1.550	1.526	Na3	2 x O	2.30(2)	2.453	2.416	0.26
P3	1 x O	1.51(2)	1.546	1.523	“1”	1 x F	2.40(2)	2.493	2.455	0.14
	1 x O	1.53(2)	1.547	1.524		2 x F	2.45(2)	2.490	2.453	0.12
	1 x O	1.53(2)	1.558	1.534		2 x O	2.50(2)	2.606	2.567	0.15
	1 x O	1.61(2)	1.540	1.517	Na4	2 x O	2.36(2)	2.394	2.358	0.22
V1	1 x O	1.92(2)	1.994	1.965	“1”	1 x F	2.40(2)	2.432	2.396	0.14
	1 x F	1.963(6)	2.025	1.995		2 x F	2.45(2)	2.458	2.421	0.13
	1 x O	1.97(1)	2.010	1.980		2 x O	2.61(2)	2.603	2.564	0.11
	1 x O	1.99(2)	2.030	2.000	Na5	2 x O	2.34(2)	2.400	2.364	0.23
	1 x O	2.02(2)	2.063	2.032	“2”	2 x F	2.52(1)	2.484	2.447	0.10
	1 x F	2.08(1)	2.031	2.000		2 x O	2.60(2)	2.653	2.613	0.12
V2	1 x F	1.90(2)	2.026	1.996		1 x F	2.62(2)	2.699	2.658	0.08
	1 x O	1.96(2)	1.991	1.961	Na6	2 x F	2.35(1)	2.379	2.343	0.16
	1 x O	1.96(2)	2.040	2.010	“3”	2 x O	2.44(2)	2.447	2.411	0.18
	1 x F	1.994(5)	2.027	1.996	Na7	2 x F	2.34(2)	2.379	2.343	0.16
	1 x O	2.02(1)	2.050	2.019	“3”	2 x O	2.43(2)	2.435	2.398	0.18
	1 x O	2.04(1)	2.018	1.988		2 x O	2.67(2)	2.804	2.762	0.09
V3	1 x O	1.92(1)	2.027	1.997	Na8	2 x F	2.45(2)	2.472	2.435	0.13
	1 x O	1.95(1)	1.998	1.968	“2”	1 x F	2.46(2)	2.492	2.455	0.12
	1 x F	1.97(2)	2.028	1.998		2 x O	2.51(2)	2.517	2.480	0.15
	1 x F	1.979(6)	2.027	1.996		2 x O	2.56(2)	2.639	2.600	0.13
	1 x O	2.01(2)	2.021	1.990	Na9	2 x F	2.43(2)	2.480	2.443	0.13
	1 x O	2.03(2)	2.070	2.039	“1”	2 x O	2.48(2)	2.469	2.432	0.16
					2 x O	2.50(2)	2.482	2.445	0.15	
					1 x F	2.51(2)	2.476	2.438	0.11	

¹Labels of “1”, “2”, and “3” indicate sites derived from the RT Na1, Na2, and Na3 sites, respectively. ²Bond valence sum (BVS) values for Na sites were calculated using the experimental XRD results and are provided on a per-bond basis to show relative bond strengths.

The LT-NVPF coordination environments for the 12 total trimers of Na sites surrounding central F ions in the optimal $A2_1am$ structure are shown in Figure 9. Due to the symmetry of this space group, there are only 3 unique types of trimers since each Na site has a multiplicity of 4. The coordination environment for each of the 9 distinct crystallographic Na sites in the LT structure can be related to one of the three types of sites in the RT polymorph. Since the RT Na1 site is already fully occupied, it is no surprise that this coordination environment is preserved in the LT polymorph. This can be seen in that each trimer of Na sites has one mono-capped trigonal prism that is derived from a Na1 site whose CN = 7 polyhedron is rendered in blue. Each trimer also contains a second Na site with a mono-capped trigonal prism that is clearly derived from a Na2 site in

the RT structure whose CN = 7 polyhedron is rendered in green. The relative orientation within the trimer of the Na2-derived site to the Na1-derived site is not always the same. Along the tripled b -axis of the unit cell, the Na1-derived sites are always oriented in the same direction while the Na2-derived sites have a repeating pattern of ++- or --+ that can only be realized within the tripled unit cell where + and - indicate the polyhedron is pointing towards the + or - a -axis, respectively).

The third type of Na site in the trimer is derived from the Na3 site in the RT polymorph, and it is the most strongly distorted relative to the parent structure. In the RT structure, the Na3 site is significantly displaced relative to the center of the trigonal prism. In the LT structure, the Na ion can be displaced even further to the point where it moves into what would have been

a face of the original trigonal prismatic environment and is better described as having a square planar environment with a stoichiometry of NaO_2F_2 . In this planar environment, the coordinating O and F anions have very short Na-X distances of about 2.4 Å, strongly contrasting with the pair of oxygens lost from the coordination shell which are pushed out to a Na-X distance of nearly 2.9 Å. For the three unique polyhedral types derived from the Na3 site of the RT structure, one has a clear square planar CN = 4 environment, another has a clear prismatic CN = 6 environment similar to that inferred from the room temperature structure, and the third is intermediate between the two.

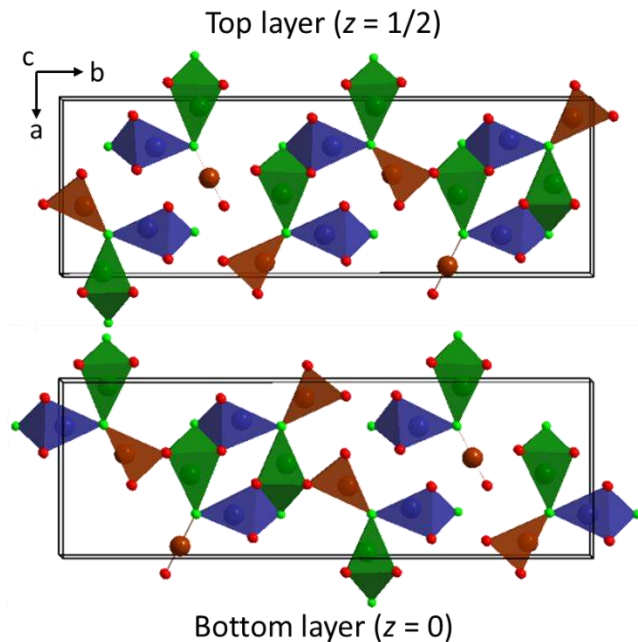


Figure 9. Coordination environments of the 36 Na atomic sites in LT-NVPF structure organized as 12 trimers (6 each at $z = 0$ and $z = 1/2$) around central fluorine sites (green spheres), allowing a comparison with other LT-NVPF structural models (Figure S6). Sites derived from Na1 (CN = 7) are shown in blue, Na2 (CN = 7) in green, and Na3 in brown (CN = 6 or 4).

Intriguingly, these types of coordination environments are robustly observed across a wide range of optimized LT-NVPF trial structures generated through enumeration followed by Rietveld refinement. A comparison of the structures returning the lowest R_{wp} value for each of the 10 space groups identified through enumeration is shown in Figure S5 and S6. In each case, it is observed for the 36 total Na sites within the unit cell volume that 12 have an environment characteristic of the Na1 site (blue), 12 have that of the Na2 site (green) and the remaining 12 have that of the Na3 site (brown) of the RT structure. While the Na1-derived site positions are the same in every refinement, the trial structures most noticeably differ in the arrangement of the Na2- and Na3-derived sites. Furthermore, the three space groups with the lowest R_{wp} fits ($P2_1nm$, $R_{wp} = 9.29$; $P2_1am$, $R_{wp} = 9.30$, $P11m$, $R_{wp} = 9.33$) adhere to the key rules for arranging Na sites observed for the optimal $A2_1am$ structure, namely that (1) inversion symmetry is not present, (2) each trimer has exactly one Na1-, Na2-, and Na3-derived site, and (3) the relative orientations of the Na2 sites repeat with a +- or -- + pattern. Interestingly, the lowest energy DFT structure identified through exhaustive evaluation of all possible orderings in the RT-NVPF subcell¹⁴ also obeyed the first two rules and has

the same space group symmetry as the optimal $A2_1am$ symmetry identified in this work (though the 3x smaller DFT cell results in a different location of some symmetry elements with respect to the polyhedral units of the structure). This space group symmetry was also reported for the doubly desodiated composition of $\text{Na}_1\text{V}_2(\text{PO}_4)_2\text{F}_3$, again with a 3x smaller cell volume.⁴ The lowest R_{wp} results for all other candidate space groups identified among the enumerated trial structures are observed to violate one or more of these rules, and for this reason they are limited in their ability to reproduce the experimentally observed diffraction pattern.

DISCUSSION

While there is no single test or analysis that can confirm the proposed $A2_1am$ symmetry structure for $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$, this structural model is consistent with many separate observations for both the refined structure and for the multitude of other trial structures. The $A2_1am$ symmetry matches the preliminary indexing of the experimental diffraction data (centered orthorhombic lattice). Indexing studies were unable to identify any simpler lattice that could generate all observed peaks nor any higher symmetry structure that could reproduce the observed diffraction data with more constraints. When DFT relaxation was carried out in the absence of symmetry constraints, the $A2_1am$ symmetry was retained. The structural features observed in the refined $A2_1am$ structure provide a clear reason for the tripling of the cell volume, and furthermore do so with a minimal number of independent crystallographic sites. Finally, there is a clear pathway of symmetry breaking connecting the proposed $A2_1am$ superstructure with the $Amam$ symmetry subcell of RT-NVPF. $A2_1am$ symmetry is one of the seven maximal non-isomorphic subgroups of the $Amam$ symmetry of the RT polymorph, and the tripled-volume $A2_1am$ supercell is the maximal isomorphic $A2_1am$ supergroup of lowest index associated with expansion in the b -axis direction.

The modeled $A2_1am$ diffraction pattern fits the observed diffraction data across hundreds of diffraction peaks with minimal errors (Figure S3). As such, if a better structural model exists, any differences with the present $A2_1am$ structure must be extremely small. In many ways, the quality of the present powder diffraction data sets the limits for the possible confidence in the correctness of the proposed $A2_1am$ structural model since the very weak nature of the superstructure peaks (~100x weaker than the subcell peaks) means that most of them have intensities which are near or below the detection limit of the high-resolution synchrotron data used for this work. The collection of complementary higher-intensity lower-resolution synchrotron powder diffraction data would enable the superstructures to be more effectively probed, and additionally, would significantly reduce the large esds of ~0.05 Å associated with key bond distances.

Even though nearly 3,000 trial structures were tested, the possible existence of a better structural description than the $A2_1am$ model that we have proposed cannot be ruled out. All the trial structures were generated under the assumption of one set of rules for ordering – it is not feasible to exhaustively examine all possible configurations under all possible rules for ordering. However, exhaustive sampling of parameter space is not required – it is only necessary to generate a trial structure which can relax to the global minimum (Fig. 4). This can occur in either the correct space group or in subgroups of the correct space group but is necessarily impossible in space groups which have symmetry constraints not present in the correct one. If different rules govern the ordering of cations, it is possible that LT-NVPF

has a different space group symmetry which was not identified through the enumeration process. However, the tested space groups effectively sampled the subgroups of the RT-NVVPF *Amam* structure. Furthermore, the low symmetry space group *P11m* provided essentially no constraints to the nature of the ordering and the 2,716 trial configurations in the space group were unable to do a substantially better job modeling the data than the much more highly constrained *A2₁am* structure. We do not believe that a better 3D space group can be found, though we cannot rule out the possibility of a higher dimensional space group providing a simpler or more complete description of the LT-NVVPF symmetry. However, an effective investigation of possible 4D, 5D, and 6D space groups would require a data set in which the different cell axes can be more effectively separated and discriminated (such as electron or single crystal diffraction data) than the present powder diffraction data.

A survey of inorganic structural databases indicates that the LT-NVVPF unit cell volume is very large, with only about 6% of structures having a larger reduced unit cell volume. Among those 6% of structural reports, the overwhelming majority of structures were solved by single crystal rather than powder diffraction methods. The few percent of structures in this tier that were solved by powder methods were generally solved by either using a closely related chemical analogue with a known crystal structure as a starting point,³⁵ or by systematically evaluating candidates identified through subgroup-supergroup relationships using distortion mode analysis techniques to minimize the number of refined parameters.³⁶ The number of unique crystallographic sites in the structure of LT-NVVPF described here is within about a factor of 4 of the largest structures robustly solved by powder diffraction for non-molecular solids, and the materials genomics method described here is thus one of the most powerful tools for powder structure determination.³⁷

CONCLUSIONS

We demonstrate a new materials genomics approach to structure solution from powder diffraction data that utilizes enumeration methods to automatically generate chemically plausible structures whose relative merit is then assessed by Rietveld refinement. This approach was used to solve the complex low-temperature structure of $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ even in the absence of conclusive information about its space group symmetry, a system that conventional structure solution methods failed to produce a structure for in prior investigations. In this materials genomics approach, the generation of structural models was predominantly driven by chemistry-based rather than symmetry-based principles, in contrast to other approaches typically used to solve very complex structures. Our novel approach should be broadly applicable for elucidating the complex ordering of cations that occur in battery materials during the processes of ion intercalation and de-intercalation that play a central role in a many important emerging energy storage systems.

ASSOCIATED CONTENT

Supporting Information.

Additional structural and fitting figures
XRD-optimized CIF file of optimal *A2₁am* structure
As-enumerated CIF file of optimal *A2₁am* structure
DFT-optimized CIF file of optimal *A2₁am* structure
DFT B-value calculation CIF file of optimal *A2₁am* structure
Synchrotron diffraction data – 100 K
Synchrotron diffraction data – 300 K
Synchrotron diffraction data – 400 K
Sample refinement control file – optimized positions

Sample refinement control file – fixed positions
CIF files of all trial refinements
Refinement output files of all trial refinements
Tabulated results summary for refined position trials
Tabulated results summary for fixed position trials

This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

* kpete@bnl.gov

Present Addresses

†Battery and Electrochemistry Laboratory, Institute of Nanotechnology, Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

Author Contributions

The project idea was conceived by PK during a sabbatical with GC. Sample synthesis and synchrotron data collection was carried out by MB with FF and under the supervision of CM, LC and ES. Enumeration was carried out by GM and JD under supervision of PK and KP. Rietveld refinement and structural analysis was carried out by GM with supervision by PK. DFT structural optimization was carried out by JD under supervision by KP, while DFT calculations of *B* values were carried out by AG with supervision by AJ. The manuscript was written through the contributions of all authors.

ACKNOWLEDGMENT

This work was intellectually led and supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division under contract no. DE-SC0012704 and DE-AC02-05CH11231 (Materials Project program KC23MP) at BNL and LBNL, respectively. The Institute Laue-Langevin (ILL) is acknowledged for funding the PhD thesis of MB. The ALBA synchrotron and the MSPD beamline are acknowledged for providing in-house beamtime. We gratefully acknowledge discussion with Nils Zimmerman (LBNL), Branton Campbell (BYU), and Alan Coelho (Coelho software) about other types of analysis approaches that were tested but were not reported here, and discussions with John Evans (Durham) about prior work on structure solution for extremely large crystal structures using powder diffraction data.

REFERENCES

- (1) Stephan, A. K., The Age of Li-Ion Batteries. *Joule* **2019**, *3* (11), 2583-2584.
- (2) Palacín, M. R.; de Guibert, A., Why do batteries fail? *Science* **2016**, *351* (6273), 1253292.
- (3) Crabtree, G., The coming electric vehicle transformation. *Science* **2019**, *366* (6464), 422-424.
- (4) Bianchini, M.; Fauth, F.; Brisset, N.; Weill, F.; Suard, E.; Masquelier, C.; Croguennec, L., Comprehensive Investigation of the $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ - $\text{NaV}_2(\text{PO}_4)_2\text{F}_3$ System by Operando High Resolution Synchrotron X-ray Diffraction. *Chem. Mater.* **2015**, *27* (8), 3009-3020.
- (5) Gover, R. K. B.; Bryan, A.; Burns, P.; Barker, J., The electrochemical insertion properties of sodium vanadium fluorophosphate, $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$. *Solid State Ionics* **2006**, *177* (17-18), 1495-1500.
- (6) Shakoar, R.; Seo, D.-H.; Kim, H.; Park, Y.-U.; Kim, J.; Kim, S.-W.; Gwon, H.; Lee, S.; Kang, K., A combined first principles and experimental study on $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ for rechargeable Na batteries. *J. Mater. Chem.* **2012**, *22* (38), 20535-20541.

- (7) Bianchini, M.; Xiao, P. H.; Wang, Y.; Ceder, G., Additional Sodium Insertion into Polyanionic Cathodes for Higher-Energy Na-Ion Batteries. *Adv. Energy Mater.* **2017**, *7* (18).
- (8) Yan, G.; Dugas, R.; Tarascon, J.-M., The $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ /Carbon Na-Ion Battery: Its Performance Understanding as Deduced from Differential Voltage Analysis. *J. Electrochem. Soc.* **2018**, *165* (2), A220-A227.
- (9) Yan, G.; Mariyappan, S.; Rousse, G.; Jacquet, Q.; Deschamps, M.; David, R.; Mirvaux, B.; Freeland, J. W.; Tarascon, J.-M., Higher energy and safer sodium ion batteries via an electrochemically made disordered $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ material. *Nat. Commun.* **2019**, *10* (1), 585.
- (10) Broux, T.; Fauth, F.; Hall, N.; Chatillon, Y.; Bianchini, M.; Bamine, T.; Leriche, J.-B.; Suard, E.; Carlier, D.; Reynier, Y.; Simonin, L.; Masquelier, C.; Croguennec, L., High Rate Performance for Carbon-Coated $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ in Na-Ion Batteries. *Small Methods* **2018**, *0* (0), 1800215.
- (11) Ponrouch, A.; Dedryvere, R.; Monti, D.; Demet, A. E.; Mba, J. M. A.; Croguennec, L.; Masquelier, C.; Johansson, P.; Palacin, M. R., Towards high energy density sodium ion batteries through electrolyte optimization. *Energy Environ. Sci.* **2013**, *6* (8), 2361-2369.
- (12) Bianchini, M.; Brisset, N.; Fauth, F.; Weill, F.; Elkaim, E.; Suard, E.; Masquelier, C.; Croguennec, L., $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$ Revisited: A High-Resolution Diffraction Study. *Chem. Mater.* **2014**, *26* (14), 4238-4247.
- (13) Bianchini, M. In situ diffraction studies of electrode materials for Li-ion and Na-ion batteries. Amiens, 2015.
- (14) Dacek, S. T.; Richards, W. D.; Kitchaev, D. A.; Ceder, G., Structure and Dynamics of Fluorophosphate Na-Ion Battery Cathodes. *Chem. Mater.* **2016**, *28* (15), 5450-5460.
- (15) Hart, G. L. W.; Forcade, R. W., Generating derivative structures from multilattices: Algorithm and application to hcp alloys. *Phys. Rev. B* **2009**, *80* (1), 014120.
- (16) Hart, G. L. W.; Nelson, L. J.; Forcade, R. W., Generating derivative structures at a fixed concentration. *Comput. Mater. Sci.* **2012**, *59*, 101-107.
- (17) Meredig, B.; Wolverton, C., A hybrid computational-experimental approach for automated crystal structure solution. *Nat. Mater.* **2013**, *12* (2), 123-127.
- (18) Fauth, F.; Peral, I.; Popescu, C.; Knapp, M., The new Material Science Powder Diffraction beamline at ALBA Synchrotron. *Powder Diffr.* **2013**, *28* (S2), S360-S370.
- (19) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G., Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **2013**, *68*, 314-319.
- (20) Hart, G. L. W.; Forcade, R. W., Algorithm for generating derivative structures. *Phys. Rev. B* **2008**, *77* (22), 224115.
- (21) Coelho, A. A., TOPAS and TOPAS-Academic: an optimization program integrating computer algebra and crystallographic objects written in C++. *J. Appl. Crystallogr.* **2018**, *51* (1), 210-218.
- (22) Kresse, G.; Furthmüller, J., Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6* (1), 15-50.
- (23) Kresse, G.; Furthmüller, J., Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54* (16), 11169.
- (24) Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865.
- (25) Mathew, K.; Montoya, J. H.; Faghaninia, A.; Dwarakanath, S.; Aykol, M.; Tang, H.; Chu, I.-h.; Smidt, T.; Bocklund, B.; Horton, M., Atomate: A high-level interface to generate, execute, and analyze computational materials science workflows. *Comput. Mater. Sci.* **2017**, *139*, 140-152.
- (26) Togo, A.; Tanaka, I., First principles phonon calculations in materials science. *Scripta Mater.* **2015**, *108*, 1-5.
- (27) Grosse-Kunstleve, R. W.; Adams, P. D., On the handling of atomic anisotropic displacement parameters. *J. Appl. Crystallogr.* **2002**, *35*, 477-480.
- (28) Gautam, G. S.; Canepa, P.; Abdellahi, A.; Urban, A.; Malik, R.; Ceder, G., The Intercalation Phase Diagram of Mg in V_2O_5 from First-Principles. *Chem. Mater.* **2015**, *27* (10), 3733-3742.
- (29) Hinuma, Y.; Meng, Y. S.; Kang, K. S.; Ceder, G., Phase transitions in the $\text{LiNi}_{0.5}\text{Mn}_{0.5}\text{O}_2$ system with temperature. *Chem. Mater.* **2007**, *19* (7), 1790-1800.
- (30) Huang, W. X.; Kitchaev, D. A.; Dacek, S. T.; Rong, Z. Q.; Urban, A.; Cao, S.; Luo, C.; Ceder, G., Finding and proving the exact ground state of a generalized Ising model by convex optimization and MAX-SAT. *Phys. Rev. B* **2016**, *94* (13).
- (31) Urban, A.; Seo, D. H.; Ceder, G., Computational understanding of Li-ion batteries. *npj Comput. Mater.* **2016**, *2*.
- (32) Van der Ven, A.; Ceder, G.; Asta, M.; Tepesch, P. D., First-principles theory of ionic diffusion with nondilute carriers. *Phys. Rev. B* **2001**, *64* (18).
- (33) Lejaeghere, K.; Van Speybroeck, V.; Van Oost, G.; Cottenier, S., Error estimates for solid-state density-functional theory predictions: an overview by means of the ground-state elemental crystals. *Crit. Rev. Solid State Mater. Sci.* **2014**, *39* (1), 1-24.
- (34) Wang, J.; Wang, Y.; Seo, D.-H.; Shi, T.; Chen, S.; Tian, Y.; Kim, H.; Ceder, G., A High-Energy NASICON-Type Cathode Material for Na-Ion Batteries. *Adv. Energy Mater.* **2020**, *10* (10), 1903968.
- (35) Stinton, G. W.; Hampson, M. R.; Evans, J. S. O., The 136-atom structure of ZrP_2O_7 and HfP_2O_7 from powder diffraction data. *Inorg. Chem.* **2006**, *45* (11), 4352-4358.
- (36) Lewis, J. W.; Payne, J. L.; Evans, I. R.; Stokes, H. T.; Campbell, B. J.; Evans, J. S. O., An Exhaustive Symmetry Approach to Structure Determination: Phase Transitions in $\text{Bi}_2\text{Sn}_2\text{O}_7$. *J. Am. Chem. Soc.* **2016**, *138* (25), 8031-8042.
- (37) Cerny, R., Crystal Structures from Powder Diffraction: Principles, Difficulties and Progress. *Crystals* **2017**, *7* (5).

