Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d6dd00064a

Data-driven exploration of AB_2X_4 (X = O, S, Se, Te) spinel chemical space

Panyalak Detrattanawichai, Zhenzhu Li, Hyunsoo Park, Kinga O. Mastej and Aron Walsh *

The discovery of materials has long been a fundamental building block for technological advancement, yet traditional trial-and-error methods are slow and costly to meet the growing demand for novel functionality, particularly in green energy technologies, energy storage, and electronics. In response to this challenge, high-throughput screening and data-driven workflows that combine computational simulations, machine learning models, and materials databases have emerged as powerful tools for accelerating materials discovery. Spinels (AB_2X_4) stand out as a versatile class of materials with applications ranging from energy storage to catalysis. However, their full compositional space remains largely unexplored. In this work, we present a data-driven framework to identify potentially synthesisable spinel compounds composed of the first 83 elements in the periodic table with oxygen (O^{2-}) and three chalcogen anions (S^{2-}, Se^{2-}, Te^{2-}). Over 30 000 charge-balanced and chemically plausible candidates, including inverse spinels, were sequentially filtered based on the stability, structural feasibility, and electronic properties criteria. Our workflow integrates materials databases, empirical heuristic rules, and machine learning predictions to efficiently reduce the candidate pool. As a result, 2303 novel spinel candidates were identified from this workflow, offering a diverse subset of target compounds for further investigation.

Received 5th February 2026
Accepted 8th April 2026

DOI: 10.1039/d6dd00064a

rsc.li/digitaldiscovery

1 Introduction

There is a transition toward clean energy solutions to tackle global challenges such as resource scarcity and climate change, both of which demand urgent and innovative solutions. As a result, the need for novel materials for green energy technology is more crucial than ever.[1–6] The discovery of materials with the necessary functionality will not only drive technological innovation, but can also support global sustainability goals.

Among many classes of materials, spinels, a family of AB_2X_4 compounds that adopt the crystal structure of the naturally occurring mineral $MgAl_2O_4$, stand out due to their structural versatility and wide-ranging properties. In the regular spinel structure, the larger X anions form a tightly packed cubic structure, with the A- and B-site cations filling the tetrahedral and octahedral sites, respectively, arranged in compliance with the space group $Fd\bar{3}m$.[7–12] A wide range of elemental substitutions at A and B sites are possible, leading to exceptional structural flexibility and a large number of naturally occurring and synthetic spinels.[13–16]

Barth and Posnjak[17] found that the A- and B-site cations can vary their distributions between the tetrahedral and octahedral voids, creating cation disordering. Spinels with A-site cations occupying tetrahedral sites and B-site cations occupying octahedral sites, as shown in Fig. 1A, are known as "normal" spinels (*e.g.* $MgAl_2O_4$). In comparison, spinels with half of the B-site cations occupying the tetrahedral sites, while the A-site cations and the other half of the B-site cations occupying the octahedral sites, as shown in Fig. 1B, are known as "inverse" spinels (*e.g.* $SnZn_2O_4$). These represent two extremes, and a parameter that describes the distribution of cations between A- and B-sites, which quantifies the likelihood of a spinel structure being "inverse", is known as the cation inversion parameter (or degree of inversion) x, ranging from 0 to 1. Therefore, the formula for AB_2X_4 spinel can be rewritten as $(A_{1-x}B_x)[A_xB_{2-x}]X_4$, where $x = 0$ refers to normal spinel and $x =$

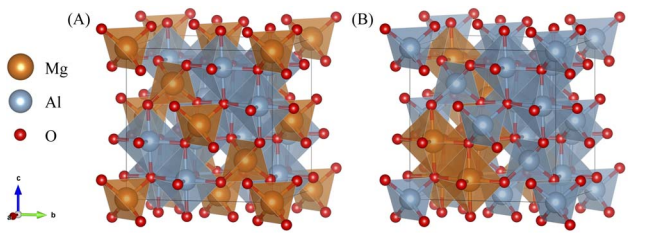

Fig. 1 (A) Normal spinel structure of $MgAl_2O_4$, where the tetrahedral and octahedral sites are occupied by Mg^{2+} and Al^{3+}, respectively. (B) Inverse spinel structure of $MgAl_2O_4$ with cation inversion parameter $x = 1$, where the tetrahedral sites are occupied by half of the Al^{3+}, while the octahedral sites are occupied by Mg^{2+} and the remaining half of the Al^{3+}.

Department of Materials, Imperial College London, London SW7 2AZ, UK. E-mail: a.walsh@imperial.ac.uk

1 refers to inverse spinel.[13,14] This distribution of cations can significantly manipulate the properties of the structure.

Disorder in the cation sublattice can impact the chemical and physical properties of spinels. This can be observed from their applications as superconductors (*e.g.* $LiTi_2O_4$),[18–20] catalysts (*e.g.* $NiAl_2O_4$),[21–24] magnetic cores (*e.g.* Fe_3O_4),[25] and gas sensors (*e.g.* $ZnFe_2O_4$).[24] With such a wide range of applications, spinel has become one of the interesting classes of compounds that could potentially be utilised in green energy technologies, such as electrochemical energy storage.[26–28]

Most spinel studies have primarily focused on naturally occurring compositions or on optimising specific properties, leaving a large number of valuable compounds potentially undiscovered.[9–12,29–32] Given the vast number of potential elemental combinations across A-, B-, and X-sites, exploring the entire compositional space of spinels through a conventional trial-and-error experimental or computational approach is impractical and unrealistic due to chemical complexity and resource demands involved. Bhattacharya and Wolverton used high-throughput first-principles calculations to explore quaternary spinels such as $LiTiFeO_4$ for Li-battery cathodes.[33] Quintero *et al.* synthesised a series of novel quaternary thiospinels, including $Li_2MgSn_3S_8$.[34] Such studies highlight the potential for realising new materials in this space.

Here, we present a data-driven workflow integrating computational simulations, machine learning models, and materials databases to efficiently explore the full compositional space of ternary spinels, focusing on compounds formed from the first 83 elements in the periodic table with four anions O^{2-}, S^{2-}, Se^{2-}, and Te^{2-}. Our workflow enables a systematic approach to narrow down the search space from over 30 000 possible compositions to 2303. These data-driven methods allow us to rapidly screen for candidates among thousands of compounds, substantially minimising the experimental and computational cost.

2 Results and discussion

2.1 Data mining

We begin by collecting experimentally reported AB_2X_4 spinel materials from the Materials Project (MP)[35] and the Inorganic Crystal Structure Database (ICSD)[36,37] to examine how known

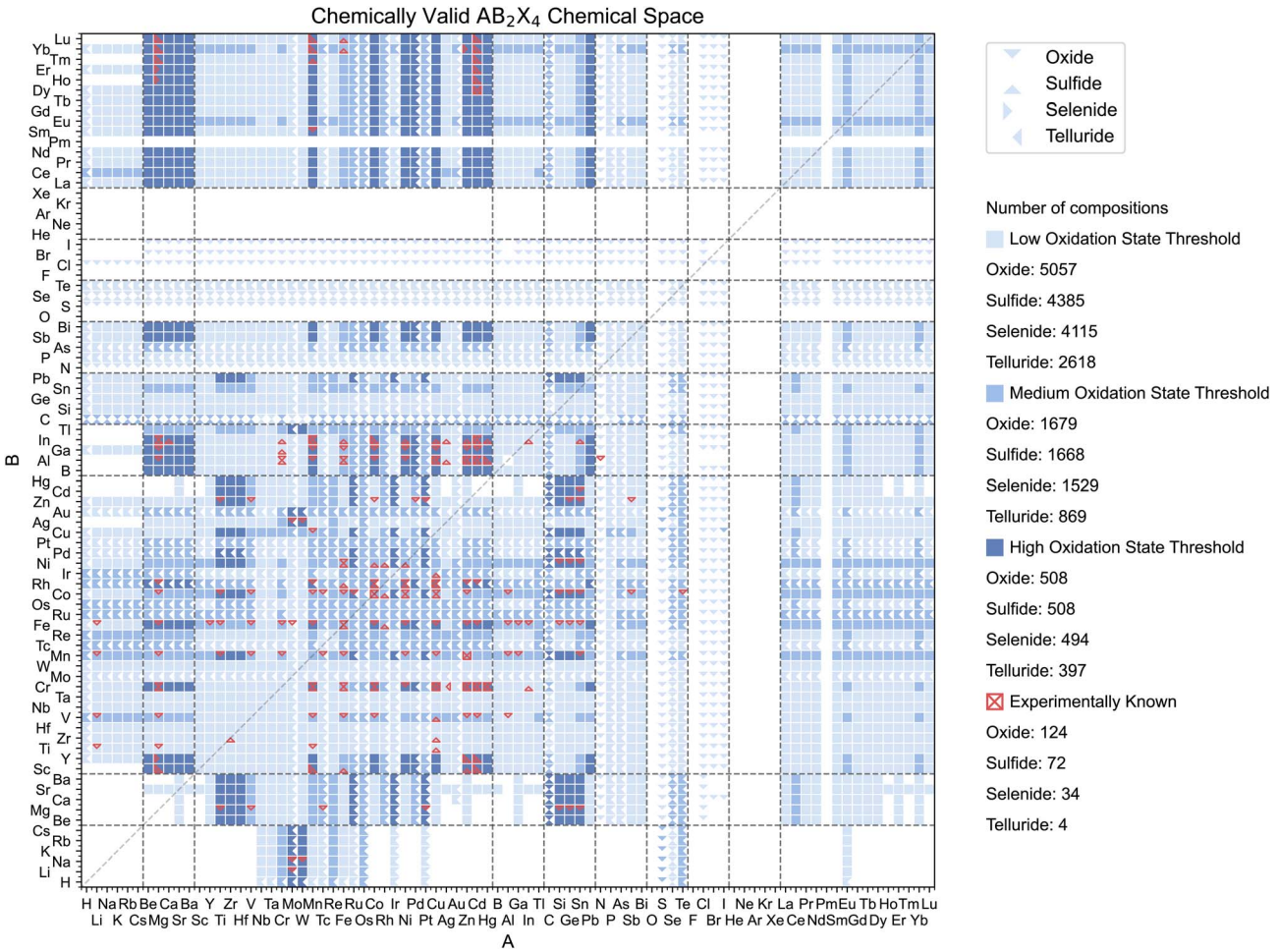

Fig. 2 Space of chemically valid AB_2X_4 spinels, generated using SMACT with three oxidation state thresholds. The intensity of blue indicates the level of oxidation state threshold. The triangle orientations represent different anion types. The red triangle markers denote the experimentally known spinel compositions. The *x*- and *y*-axes correspond to A-site and B-site cations, respectively, in the order of periodic table group (from group IA to the lanthanides). The diagonal line denote the A_3X_4 spinels where the same cation occupies both A and B sites.

spinels are distributed over the chemical space. For the MP, the MP API is employed to retrieve entries with the formulas AB_2X_4 and A_3X_4 in the $Fd\bar{3}m$ space group. For ICSD, a manual search is carried out for entries with space group $Fd\bar{3}m$ and structural type similar to $MgAl_2O_4$. The retrieved entries are then filtered to retain only compositions matching the AB_2X_4 or A_3X_4 formula. To ensure consistency, compositions with non-integer stoichiometric coefficients are rounded to the nearest integers.

Combining both datasets, we identify 234 known spinels in this space. Among them, there are 124 oxides, 72 sulfides, 34 selenides, and 4 tellurides. These known compositions are highlighted in Fig. 2 using red triangle markers, where different triangle orientations indicate different anion types.

2.2 Generating spinel compositions

We start by applying a series of chemical rules-including charge neutrality and electronegativity balance-to efficiently screen for chemically plausible compositions composed of the first 83 elements, excluding noble gases, in the periodic table with oxygen and three chalcogen anions. These rules are implemented in the Semiconducting Materials by Analogy and Chemical Theory (SMACT)[38,39] package.

Spinels are heteropolar compounds and can be classified based on how the charges are balanced. For the general formula AB_2X_4, charge neutrality requires that the sum of cationic charge satisfies $q^A + 2q^B = -4q^X$. For oxides and chalcogenides, where $q^X = -2$, this leads to three common charge combinations: ($q^A = 2$, $q^B = 3$), ($q^A = 4$, $q^B = 2$), and ($q^A = 6$, $q^B = 1$), with known examples being $MgAl_2O_4$, $TiMg_2O_4$, and WLi_2O_4, respectively.

To note here, the oxidation states of elements used during the SMACT screening include every possible oxidation state that are found in ICSD data, without considering the imbalanced distribution of commonly or uncommonly seen oxidation states. In response to this, three oxidation state thresholds-low, medium, and high-to filter for chemically valid compositions are implemented. The low threshold includes all reported oxidation states that appear at least 3 times; the medium threshold excludes those that appear less than 10% of total appearance for a given element; and the high threshold retains only the most common oxidation state for each element.

The resulting chemically valid compositions are represented in Fig. 2, where different triangle orientations indicate different anion type (same as the known ones), and the intensity of the blue represents the three levels of oxidation state thresholds. Using SMACT, we obtain 16 175, 5745, and 1907 valid compositions for the low-, medium-, and high-threshold filters. "Recover" is used here to refer to the fraction of experimentally known spinels that are successfully preserved after applying the screening filter. In detail, the low threshold recovers 100% of the known AB_2X_4 spinels. However, the number of recovered known spinels drops by 9% at the medium threshold and by 40% at the high threshold. This highlights the importance of oxidation state selection when identifying potential candidates. To balance chemical plausibility with compositional flexibility-while ignoring rare oxidation states-the medium threshold is adopted for the remainder of our investigation.

2.3 Structural optimisation

To proceed with further calculations, we first generate two prototype structures with lattice parameter $a = 8$ Å: one representing the normal spinel configuration, resembling that of $MgAl_2O_4$, and another representing the inverse configuration, generated using the Special Quasirandom Structure (SQS)[40,41] approach based on the normal prototype. These prototype structures serve as initial structures for all compositions generated using SMACT. Therefore, we have more than 10 000 structures to optimise. It is important to note that the normal and inverse structures for A_3X_4 spinels are the same since they only possess one type of cation. To efficiently screen the large number of candidate structures, a machine learning force-field model pre-trained with inorganic materials, the Multiple Atomic Cluster Expansion (MACE),[42–44] is employed to accelerate the optimisation of crystal structures. The foundation model MACE-MP-0a is used to perform structural relaxation *via* the Atomic Simulation Environment (ASE) using the FIRE structural optimisation algorithm. Both atomic positions and cell are allowed to relax freely through the Frechet cell filter with a loose force convergence criterion $f_{max} = 0.05$ eV Å^{-1}. To avoid excessive calculations on structures that fail to converge, the maximum number of optimisation steps is set to 600. In the calculations, any structure that fails to converge within this limit or possesses excessively large forces is discarded from further analysis.

To verify the accuracy of the MACE-MP-0a for the spinels family, comparative calculations on a representative set of the spinel chemical space (10 oxides, 10 sulfides, 10 selenides, and 4 tellurides) have been carried out using MACE-MP-0a and conventional Density Functional Theory (DFT)[45,46] with Perdew–Burke–Ernzerhof (PBE)[47] functional within Vienna *Ab initio* Simulation Package (VASP).[48,49] The resulting parity and difference plots for lattice parameters and total energy per unit cell are provided in the SI as Fig. S1 and S2 with the computational details. The results show that while MACE-MP-0a is powerful for oxides, an increase in uncertainty of prediction is observed for the chalcogenides, especially tellurides, which likely reflects the decreasing abundance of relevant training data for these chemistries in the MACE-MP-0a model.

2.4 Crystal likelihood

After obtaining the optimised structures and their total energies, we calculate the energy above the convex hull (E_{hull}) for each structure based on the MACE-predicted energy using functions from Python Materials Genomics (Pymatgen).[50] Notably, Pymatgen phase diagram module could not calculate E_{hull} of Yb due to data deprecated on MP. Therefore, compositions with Yb are excluded from the screening. For compositions where both the normal and inverse configurations successfully converged, the structures with the lower E_{hull} are selected for further analysis. This method allows for evaluating the crystal likelihood using thermodynamic approach for each structure and visualise the distribution of E_{hull} across different anion types, as shown in Fig. S3 in the SI. While $E_{hull} = 0$ corresponds to the prediction of a stable compound on the

convex hull from the MP database, a metastability window exists where materials have been experimentally realised. In this work, we define a metastability window of $E_{\text{hull}} \leq 0.2$ eV per atom.

Beyond pure thermodynamic considerations, data-driven measures have also been developed to assess how likely it is that a new compound can be synthesised.[51–56] To further refine the candidate screening, a data-driven machine learning model, namely, synthesisability-stoi-CGNF, based on Positive Unlabelled Learning (PUL) proposed by Jang *et al.*[53] is also adopted to investigate the crystal likelihood from a different perspective. This model is trained using stoichiometric crystal data, where known compositions are labelled as 1 (positive) and unknown as 0. The output of this model is a metric, namely, crystal-likeness score (CLscore), ranging from 0 to 1, where higher values indicate an increased likelihood of compound stability. According to Jang's study, compositions with CLscore ≥ 0.5 are considered potentially stable. However, since this model only takes stoichiometry (or compositions) as input, it cannot distinguish the difference between normal and inverse spinels. Nevertheless, this method offers us a complementary perspective on the crystal likelihood of spinels based on existing materials data. The distribution of CLscore is illustrated in Fig. S4 in the SI.

$$N_{E_{\text{hull}}} = 1 - \frac{E_{\text{hull}}}{\alpha} \tag{1}$$

$$N_{\text{CLscore}} = \frac{\text{CLscore} - \beta}{\beta} \tag{2}$$

$$S_{\text{score}} = \frac{N_{E_{\text{hull}}} + N_{\text{CLscore}}}{2} \tag{3}$$

From the thermodynamic and data-driven evaluations, we combine the results by choosing only structures that satisfy both criteria ($E_{\text{hull}} \leq 0.2$ eV per atom and CLscore ≥ 0.5). Each metric is first normalised using eqn (1) and (2), where α and β refer to a metastability window of E_{hull} and stability threshold of CLscore, respectively. For example, $\alpha = 0.2$ eV per atom and $\beta = 0.5$ in Fig. 3. The normalised values are then averaged to define a new single unified metric called super score (S_{score}) as shown in eqn (3). This score integrates both E_{hull} and CLscore to quantify the joint likelihood of each composition on a scale from 0 (unlikely) to 1 (likely). The chemical space of oxide spinels with their S_{score} is depicted in Fig. 3A, with the distribution shown in Fig. S5 in the SI. This metric allows us to rank and prioritise spinel candidates. The chemical space of other types of spinel (X = S, Se, and Te) with their S_{score} are shown in Fig. S6–S8 in the SI.

The recovered percentage of the experimentally known oxide spinels under the medium oxidation state threshold at different E_{hull}, CLscore, and S_{score} can be illustrated as Fig. 4 (The same illustrations for sulfides, selenides, and tellurides are provided in the SI Fig. S9). The initial recovered percentage starts at

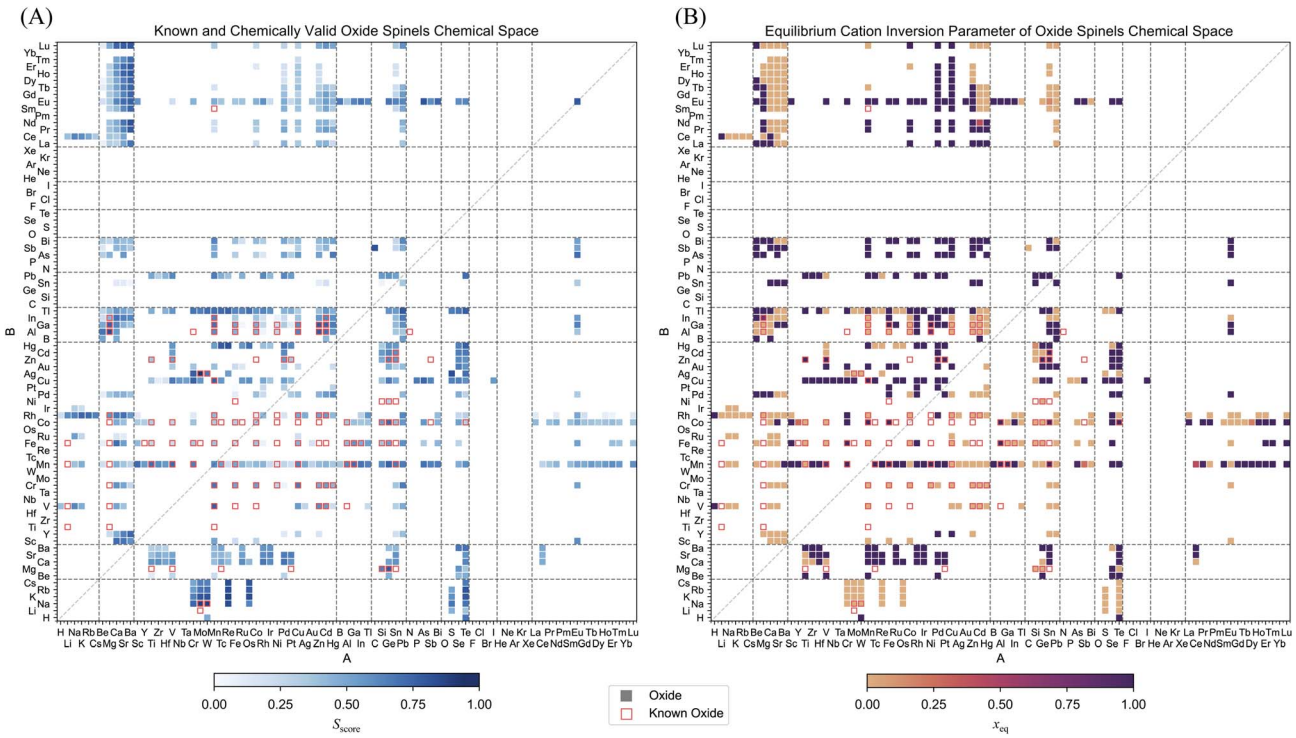

Fig. 3 Chemical space of oxide spinels that satisfy both $E_{\text{hull}} \leq 0.2$ eV per atom and CLscore ≥ 0.5. (A) The blue gradient indicates S_{score}, where lighter and darker shades designate lower and higher crystal likelihood, respectively. (B) The colourmap represents the equilibrium cation inversion parameter (x_{eq}), where lighter orange designates the normal configuration, and darker purple represents the inverse configuration. The red square markers denote the experimentally known spinel compositions.

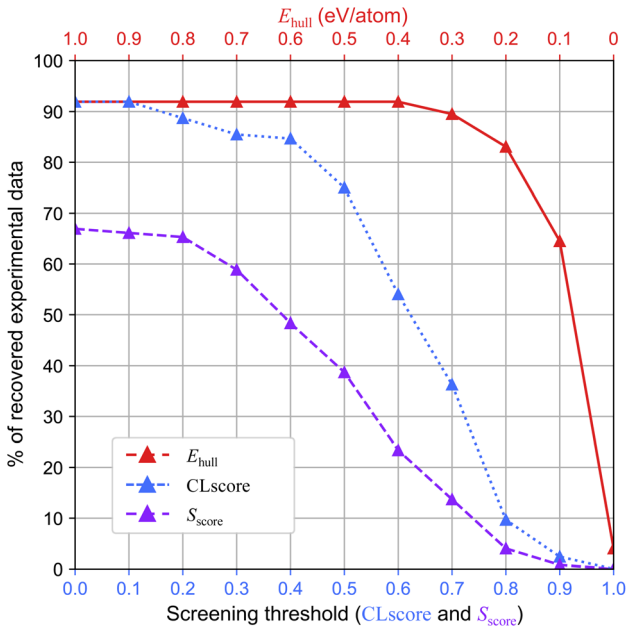

Fig. 4 Fraction of recovered experimentally known oxide spinels with medium oxidation state threshold for the initial chemical filter.

91.9%, reflecting that 8.1% of known compositions are excluded by the initial chemical filter (SMACT) or failed to converge during structural optimisation using the MACE force field. From this baseline, stricter screening filter naturally reduces the recovery rate further.

For the thermodynamic filter (E_{hull}), which is represented by the red solid line, the recovery rate remains constant above E_{hull} = 0.4 eV per atom but decreases to 83.1% at the metastability window of E_{hull} = 0.2 eV per atom. This indicates that the thermodynamic filter successfully captures the majority of the known oxide spinels with 83.1% true positive rate. The data-driven filter (CLscore) represented by the blue dotted line, on the other hand, is more selective, recovering 75% of the known oxide spinels at the stability criterion of CLscore = 0.5. This shows slightly weaker performance than the thermodynamic approach, yet it still provides a reasonable prediction of the crystal likelihood based on existing data. These results support our chosen criteria for identifying crystal likelihood as they preserve a large fraction of the known oxide spinel compositions. With both filters combined into S_{score}, which is represented by the purple dashed line, the recovery rate becomes 66.9% at S_{score} = 0. It is important to note that this 66.9% recovery rate does not represent a failure to identify valid materials, but rather the strictness of our combined screening filters that prioritises high-confidence predictions. The excluded 33.1% largely consist of rare elements where the underlying training data for the ML model is scarce. This suggests that S_{score} offers a continuous and balanced metric for evaluating crystal likelihood.

From Fig. 2 and 3A, it can be seen that a large portion of the chemical space has been filtered out during the screening process, retaining only compositions that pass through chemical filter and MACE structural optimisations, while satisfying both the thermodynamic and data-driven stability criteria. This significant reduction reflects the strictness of the combined filters and indicates that many compositions are unlikely to be synthesisable. However, it is important to note that this may not always reflect true instability since some valid but rare compositions may be excluded due to limitations in the force field or machine learning models used in the workflow. This point is supported by the fact that 66.9% of the known oxide spinels pass through the thermodynamic and data-driven filters, where some of the known oxide spinels that are chemically valid but contain rare elements, such as $TcCo_2O_4$, $NiRh_2O_4$, and $MnSm_2O_4$, are filtered out. This shows that data scarcity of compositions containing such elements can impact model predictions and influence the screening.

Moreover, it is shown in Fig. 3A that compositions with high S_{score} are spatially clustered in the chemical space. This clustering suggests that compositions with similar chemical structure tend to share similar crystal likelihood profiles. These patterns can be used to guide exploration regions in the chemical space that may be particularly promising for further studies.

2.5 Cation inversion

We have not yet distinguished between normal and inverse spinels. $MgAl_2O_4$ is a normal spinel if naturally grown ($x \approx 0$). However, when exposed to a high temperature (*ca.* 900 °C), the cations exhibit disordering with $x \approx 0.3$.[57] Furthermore, it is found that synthetic $MgAl_2O_4$ usually exhibit cation disordering with $x \sim 0.1$ to 0.6.[13] The disorder has been analysed and modelled using methods such as X-ray diffraction,[58] neutron diffraction,[9,59] Mössbauer spectroscopy,[9,60] classical electrostatics,[61] and ligand field theory.[62]

$$x_{eq} = \frac{\sqrt{9 + 8(c-1)} - 3}{2(c-1)}, \quad (4)$$

One driving force for cation inversion is configurational entropy. The normal spinel corresponds to a single configuration of atoms in a unit cell, while the inverse spinel features a range of possible distributions, which can lower the free energy of the compound. The role of entropy becomes more important at higher temperatures. Wei and Zhang[14] proposed a thermodynamic model based on the earlier work of Navrotsky and Kleppa[63] to estimate the equilibrium cation inversion parameter (x_{eq}) at a specific temperature using the energy difference between the normal and inverse configuration, as expressed in eqn (4), where $c = e^{\Delta E/kT}$, $\Delta E = E_{inverse} - E_{normal}$, k is Boltzmann constant, and T is temperature. Using this model, they successfully predicted the equilibrium cation inversion parameter for several oxide spinels. We adopt their approach to cation inversion at room temperature (298.15 K) across the spinel chemical space using MACE-predicted energies for both normal and inverse structures.

The results for oxide spinels are illustrated in Fig. 3B, and the results for the rest are provided as SI. This approach cannot describe inversion of A_3X_4 spinels, as the current force fields

don't distinguish between different species (charges) of the same element. Therefore, A_3X_4 spinels are excluded from Fig. 3B. Interestingly, consistent with what was observed using S_{score} in Fig. 3A, the equilibrium cation inversion parameter shows clear clustering of normal and inverse configurations across the chemical space. Compositions with similar chemical features exhibit similar cation disordering behaviour.

To determine the preferred configuration, we compare the E_{hull} of the normal and inverse configurations and select the one with lower E_{hull} for each composition. In cases where both configurations possess the same E_{hull}, configuration with lower total energy is chosen. We further validated these preferences by mapping them onto orbital radii maps[64,65] and comparing them against preferences determined by x_{eq}, which are rounded to 0 and 1. The resulting maps as shown in Fig. S10 and S11 show identical distributions, demonstrating that the energetically more stable structure, which is indicated by the lower E_{hull}, aligns with the calculated x_{eq}. Therefore, we adopt the lower E_{hull} configuration as the representative structure for each composition.

2.6 Candidate spinels

The number of candidate compositions remaining after each step of the screening process are summarised in Table 1. The pool of possible compositions for each anion is $83^2 = 6889$. By applying the chemical filter using SMACT with the medium oxidation state threshold, the number is reduced by over 75% for oxide, sulfide, and selenide spinels, and even more by 87.3% for telluride spinel. This shows the critical role of basic chemical rules in materials screening. By applying thermodynamic filter ($E_{hull} \leq 0.2$ eV per atom), the pool is further reduced by an average of 11.15% of the possible compositions. On the other hand, the data-driven filter (CLscore $\geq$ 0.5) is slightly less selective by narrowing down an average of 8.2%. However, with the two filters combined, the pool is effectively reduced to an average of 8.4% of the possible compositions. This highlights the utility of unified metrics that combine thermodynamic and data-driven statistical information.

Based on S_{score}, top candidates for each type of spinel are identified by ranking all compositions. This represents the raw output of the models that has not been human-filtered. Table 2 shows the top 10 spinel candidates for oxides, sulfides, selenides, and tellurides, excluding compositions that are presented on MP and ICSD, and it is noteworthy to mention that all 40 candidates favor normal configuration based on the E_{hull} value. Each candidate was checked by a manual literature search to confirm that they have not yet been experimentally reported.

The ranking procedure produced some unexpected candidates, such as SAg_2O_4. This is stoichiometrically equivalent to silver sulfate (Ag_2SO_4), whose known ground-state structure is orthorhombic, not the predicted cubic spinel configuration. To further evaluate the dynamic stability of the surviving candidates, phonon dispersion calculations were carried out through the MatCalc[66] package using the MACE-MP-0a model. The results are embedded in Table 2, where compositions shown in bold possess no imaginary phonon modes, indicating dynamic stability with respect to small atomic displacements. Full phonon dispersion plots for each candidate that possess no imaginary modes are provided in the SI, Fig. S12–S23.

Six of the oxide candidates were found to be dynamically stable. Among these, the alkali rhenates ($ReNa_2O_4$, ReK_2O_4, $ReRb_2O_4$) and $BaNd_2O_4$ are unusual due to their large B-site. The large B-site channels suggest potential utility in electrochemical applications, where such open frameworks often facilitate rapid ionic diffusion for battery electrolytes. Their

Table 1 Number of candidate AB_2X_4 compositions after each step of screening starting from a pool of 83 elements for the A and B sites

Type of anions	Known	Possible	Chemical filter	$E_{hull} \leq 0.2$ eV per atom	CLscore ≥ 0.5	$E_{hull} \leq 0.2$ eV per atom and CLscore ≥ 0.5
Oxide	124	6889	1679	885	1096	704
Sulfide	72	6889	1668	736	968	675
Selenide	34	6889	1529	671	1033	634
Telluride	4	6889	869	370	394	290

Table 2 Top 10 hypothetical spinel candidates from the combined thermodynamic and statistical S_{score} metric. Compositions shown in bold were found to possess no imaginary mode in their phonon dispersions, indicating they are stable local minima in their respective potential energy surfaces. n- and i-AB_2X_4 refer to normal and inverse configurations, respectively

Type of anions	Compositions				
Oxide	**n-ReK_2O_4**	**n-OsK_2O_4**	**n-$ReNa_2O_4$**	n-SAg_2O_4	n-$ReCs_2O_4$
	n-$ReRb_2O_4$	n-$ReHg_2O_4$	n-CSb_2O_4	**n-$BaNd_2O_4$**	**n-WK_2O_4**
Sulfide	n-CTl_2S_4	n-Pd_3S_4	n-$PbGa_2S_4$	n-$ReRb_2S_4$	n-CSn_2S_4
	n-$PbTl_2S_4$	n-$OsTb_2S_4$	n-$EuBi_2S_4$	n-WK_2S_4	n-$CoTb_2S_4$
Selenide	n-$TeCs_2Se_4$	n-SrB_2Se_4	n-$CeTl_2Se_4$	**n-Pd_3Se_4**	n-EuB_2Se_4
	n-$CaSb_2Se_4$	n-CaB_2Se_4	n-$CaTl_2Se_4$	n-Rh_3Se_4	n-$CsRh_2Se_4$
Telluride	**n-$MgNd_2Te_4$**	n-$ZrBa_2Te_4$	**n-$MgEr_2Te_4$**	**n-$MgTb_2Te_4$**	n-$MgSm_2Te_4$
	n-$MgGd_2Te_4$	**n-$MgLu_2Te_4$**	n-MgY_2Te_4	n-$CeBa_2Te_4$	n-$CdGd_2Te_4$

predicted stability here indicates they are valid local minima on the potential energy surface. One of the most plausible candidates is WK_2O_4, a finding supported by the fact that its analogue, WNa_2O_4, is known to adopt a high-pressure spinel phase.[67] Beyond structural feasibility, the WNa_2O_4 spinel possesses a wide band gap of 4.4 eV,[68] making it suitable for applications in electrocatalysis or as optoelectronic materials. We therefore anticipate WK_2O_4 will exhibit a similarly wide band gap.

In contrast, no dynamically stable sulfide spinels were identified. This may be a false negative resulting from the under-representation of ternary chalcogenides within the model's training data relative to oxides. For the selenides, Pd_3Se_4 is predicted to be stable. This composition implies a mixed-valence state of $[Pd^{2+}][Pd^{3+}]_2Se_4$. Its primary competing phase in the Pd–Se system would be the mineral verbeekite ($PdSe_2$), which contains Pd^{2+} cations charge-balanced by diselenide dimers, $(Se_2)^{2-}$. Finally, five telluride compositions show no imaginary modes, all of which conform to the $Mg(RE)_2Te_4$ formula (RE = Nd, Gd, Lu, Er, Tb). The incorporation of these magnetic rare-earth elements makes these compounds interesting targets for novel electronic and magnetic applications.

The dynamical stabilities of selected MACE force field predicted stable structures were further tested using DFT PBE functional on a $2 \times 2 \times 2$ supercell. The analysis of phonon dispersion was carried out using Phonopy package[69,70] comparing against MACE-predicted results, the comparisons are illustrated as Fig. S24–S26 in the SI with the computational details. The comparison demonstrates good agreement between DFT-PBE and MACE-MP-0a, confirming the reliability of the ML model. However, for the case of WK_2O_4, even though MACE predicted full stability, the DFT reveals a small imaginary mode at the *L*-point. This does not confirm that WK_2O_4 is unstable. These imaginary modes at zone boundaries often indicate a subtle structural distortion that may occur due to the structure falling to the local minima or a slight symmetry breaking that is energetically close to the ideal cubic phase. Such modes are often stabilised at finite temperatures, suggesting that the ideal cubic phase could still be accessible under specific growth conditions.

3 Conclusion

We have showcased a data-driven workflow to identify potentially synthesisable AB_2X_4 spinel oxides and chalcogenides across the vast chemical space composed of the first 83 elements. By integrating chemical plausibility checks, efficient structural optimisation, and stability assessments through both thermodynamic filtering and data-driven approaches, 55 112 possible compositions, including inverse spinels, are narrowed down to just 2303 compositions. We introduce a unified metric combining thermodynamics (E_{hull}) with the statistical CLscore and are able to quantify the crystal likelihood of compositions. The screening recovers over two-thirds of the known oxide spinels using a metastability window of $E_{hull} \leq 0.2$ eV per atom and potentially stable criteria of CLscore ≥ 0.5, while also uncovering hundreds of previously unknown oxide compositions. Furthermore, our investigation into the equilibrium cation inversion parameter reveals trends in the chemical space, where similar chemical structures possess similar cation disordering behaviour. The final top candidates for each anion type highlight a set of promising novel candidates. Among the top candidates, six promising candidates without imaginary phonon modes are found for oxides, none for sulfides, one for selenides, and five for tellurides through phonon dispersion calculations using MACE-MP-0a. Overall, this framework serves as a scalable workflow for materials exploration across vast chemical spaces, offering a foundation for further modelling and future experimental validations.

Author contributions

The author contributions have been defined following the CRediT system. P. D.: conceptualisation, investigation, formal analysis, methodology, visualisation, writing – original draft. Z. L.: methodology, supervision, writing – review & editing. H. P.: methodology, supervision, writing – review & editing. K. O. M.: methodology, software, writing – review & editing. A. W.: conceptualisation, methodology, supervision, writing – review & editing.

Conflicts of interest

There are no conflicts to declare.

Data availability

The workflows employ a range of open-source software packages that are cited in the main text. The original workflows and data will be made available in a GitHub repository on https://github.com/WMD-group/Spinel-chemical-space.

Supplementary information (SI) is available. See DOI: https://doi.org/10.1039/d6dd00064a.

Acknowledgements

We thank Anthony Onwuli for useful discussions and guidance on SMACT. This work was supported by the Development and Promotion of Science and Technology Talents Project (Thai Government Scholarship). *Via* our membership of the UK's HEC Materials Chemistry Consortium, which is funded by EPSRC (EP/X035859/1), this work used the ARCHER2 UK National Supercomputing Service (https://www.archer2.ac.uk). We are grateful to the UK Materials and Molecular Modelling Hub for computational resources, which is partially funded by EPSRC (EP/T022213/1, EP/W032260/1 and EP/P020194/1).

Notes and references

1 A. R. Oganov, C. J. Pickard, Q. Zhu and R. J. Needs, *Nat. Rev. Mater.*, 2019, **4**, 331–348.
2 J. N. Cawse, *Experimental design for combinatorial and high throughput materials development*, Citeseer, 2003.

3 A. Valero, A. Valero, G. Calvo and A. Ortego, *Renew. Sustain. Energy Rev.*, 2018, **93**, 178–200.
4 Y. Liu, T. Zhao, W. Ju and S. Shi, *J. Materiomics*, 2017, **3**, 159–177.
5 J. E. Saal, A. O. Oliynyk and B. Meredig, *Annu. Rev. Mater. Res.*, 2020, **50**, 49–69.
6 P. Raccuglia, K. C. Elbert, P. D. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier and A. J. Norquist, *Nature*, 2016, **533**, 73–76.
7 W. H. Bragg, *London, Edinburgh Dublin Philos. Mag. J. Sci.*, 1915, **30**, 305–315.
8 S. Nishikawa, *Proc. Tokyo Math.-Phys. Soc.*, 1915, **8**, 199–209.
9 K. R. Sanchez-Lievanos, J. L. Stair and K. E. Knowles, *Inorg. Chem.*, 2021, **60**, 4291–4305.
10 B. Xu and S. Meng, *J. Power Sources*, 2010, **195**, 4971–4976.
11 Y. Zhou, S. Sun, C. Wei, Y. Sun, P. Xi, Z. Feng and Z. J. Xu, *Adv. Mater.*, 2019, **31**, 1902509.
12 M. Nakayama, H. Taki, T. Nakamura, S. Tokuda, R. Jalem and T. Kasuga, *J. Phys. Chem. C*, 2014, **118**, 27245–27251.
13 K. E. Sickafus and R. Hughes, *J. Am. Ceram. Soc.*, 1999, **82**, 3277–3278.
14 S.-H. Wei and S. B. Zhang, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2001, **63**, 045112.
15 R. J. Hill, J. R. Craig and G. Gibbs, *Phys. Chem. Miner.*, 1979, **4**, 317–339.
16 C. Biagioni and M. Pasero, *Am. Mineral.*, 2014, **99**, 1254–1264.
17 T. F. Barth and E. Posnjak, *Z. Kristallogr. - Cryst. Mater.*, 1932, **82**, 325–341.
18 D. Johnston, H. Prakash, W. Zachariasen and R. Viswanathan, *Mater. Res. Bull.*, 1973, **8**, 777–784.
19 D. Johnston, *J. Low Temp. Phys.*, 1976, **25**, 145–175.
20 R. McCallum, D. Johnston, C. Luengo and M. Maple, *J. Low Temp. Phys.*, 1976, **25**, 177–193.
21 N. Grimes, *Phys. Technol.*, 1975, **6**, 22.
22 H. Qin, Y. He, P. Xu, D. Huang, Z. Wang, H. Wang, Z. Wang, Y. Zhao, Q. Tian and C. Wang, *Adv. Colloid Interface Sci.*, 2021, **294**, 102486.
23 J. H. Advani, G. S. More and R. Srivastava, *Green Chem.*, 2022, **24**, 3574–3604.
24 S. J. Salih and W. M. Mahmood, *Heliyon*, 2023, **9**, year.
25 M. Kurian and S. Thankachan, *Open Ceram.*, 2021, **8**, 100179.
26 D. Liu, W. Zhu, J. Trottier, C. Gagnon, F. Barray, A. Guerfi, A. Mauger, H. Groult, C. Julien, J. Goodenough, *et al.*, *RSC Adv.*, 2014, **4**, 154–167.
27 J. Bhattacharya and C. Wolverton, *Phys. Chem. Chem. Phys.*, 2013, **15**, 6486–6498.
28 T. Nissinen, T. Valo, M. Gasik, J. Rantanen and M. Lampinen, *J. Power Sources*, 2002, **106**, 109–115.
29 I. Ganesh, *Int. Mater. Rev.*, 2013, **58**, 63–112.
30 M. Liu, Z. Rong, R. Malik, P. Canepa, A. Jain, G. Ceder and K. A. Persson, *Energy Environ. Sci.*, 2015, **8**, 964–974.
31 V. Kocevski, G. Pilania and B. P. Uberuaga, *J. Mater. Chem. A*, 2020, **8**, 25756–25767.
32 F. Gao, X. Tang, H. Yi, S. Zhao, W. Zhu and Y. Shi, *J. Environ. Sci.*, 2020, **89**, 145–155.
33 J. Bhattacharya and C. Wolverton, *J. Electrochem. Soc.*, 2014, **161**, A1440.
34 M. A. Quintero, S. Hao, S. V. Patel, J.-K. Bao, X. Zhou, Y.-Y. Hu, C. Wolverton and M. G. Kanatzidis, *Chem. Mater.*, 2021, **33**, 2080–2089.
35 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, *APL Mater.*, 2013, **1**, 011002.
36 G. Bergerhoff, I. Brown, F. Allen, *et al.*, *International Union of Crystallography*, Chester, 1987, vol. 360, pp. 77–95.
37 D. Zagorac, H. Müller, S. Ruehl, J. Zagorac and S. Rehme, *J. Appl. Crystallogr.*, 2019, **52**, 918–925.
38 D. W. Davies, K. T. Butler, A. J. Jackson, A. Morris, J. M. Frost, J. M. Skelton and A. Walsh, *Chem*, 2016, **1**, 617–627.
39 D. W. Davies, K. T. Butler, A. J. Jackson, J. M. Skelton, K. Morita and A. Walsh, *J. Open Source Softw.*, 2019, **4**, 1361.
40 A. Zunger, S.-H. Wei, L. G. Ferreira and J. E. Bernard, *Phys. Rev. Lett.*, 1990, **65**, 353.
41 A. van de Walle, P. Tiwary, M. de Jong, D. L. Olmsted, M. Asta, A. Dick, D. Shin, Y. Wang, L.-Q. Chen and Z.-K. Liu, *Calphad*, 2013, **42**, 13–18.
42 I. Batatia, D. P. Kovacs, G. N. C. Simm, C. Ortner and G. Csanyi, *Adv. Neural Inf. Process. Syst.*, 2022, 11423–11436.
43 I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovács, J. Riebesell, X. R. Advincula, M. Asta, M. Avaylon, W. J. Baldwin, F. Berger, N. Bernstein, A. Bhowmik, S. M. Blau, V. Cărare, J. P. Darby, S. De, F. D. Pia, V. L. Deringer, R. Elijošius, Z. El-Machachi, F. Falcioni, E. Fako, A. C. Ferrari, A. Genreith-Schriever, J. George, R. E. A. Goodall, C. P. Grey, P. Grigorev, S. Han, W. Handley, H. H. Heenen, K. Hermansson, C. Holm, J. Jaafar, S. Hofmann, K. S. Jakob, H. Jung, V. Kapil, A. D. Kaplan, N. Karimitari, J. R. Kermode, N. Kroupa, J. Kullgren, M. C. Kuner, D. Kuryla, G. Liepuoniute, J. T. Margraf, I.-B. Magdău, A. Michaelides, J. H. Moore, A. A. Naik, S. P. Niblett, S. W. Norwood, N. O'Neill, C. Ortner, K. A. Persson, K. Reuter, A. S. Rosen, L. L. Schaaf, C. Schran, B. X. Shi, E. Sivonxay, T. K. Stenczel, V. Svahn, C. Sutton, T. D. Swinburne, J. Tilly, C. van der Oord, E. Varga-Umbrich, T. Vegge, M. Vondrák, Y. Wang, W. C. Witt, F. Zills and G. Csányi, *J. Chem. Phys.*, 2025, **163**, 184110.
44 I. Batatia, S. Batzner, D. P. Kovács, A. Musaelian, G. N. Simm, R. Drautz, C. Ortner, B. Kozinsky and G. Csányi, *Nat. Mach. Intell.*, 2025, **7**, 56–67.
45 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864.
46 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133.
47 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.
48 G. Kresse and J. Furthmüller, *Comput. Mater. Sci.*, 1996, **6**, 15–50.
49 G. Kresse and J. Furthmüller, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 1996, **54**, 11169–11186.
50 S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson and G. Ceder, *Comput. Mater. Sci.*, 2013, **68**, 314–319.

51 N. C. Frey, J. Wang, G. I. Vega Bellido, B. Anasori, Y. Gogotsi and V. B. Shenoy, *ACS Nano*, 2019, **13**, 3031–3041.
52 G. H. Gu, J. Jang, J. Noh, A. Walsh and Y. Jung, *npj Comput. Mater.*, 2022, **8**, 71.
53 J. Jang, J. Noh, L. Zhou, G. H. Gu, J. M. Gregoire and Y. Jung, *Matter*, 2024, **7**, 2294–2312.
54 S. Amariamir, J. George and P. Benner, *Digital Discovery*, 2025, **4**, 1437–1448.
55 S. Kim, Y. Jung and J. Schrier, *J. Am. Chem. Soc.*, 2024, **146**, 19654–19659.
56 V. Chung, A. Walsh and D. J. Payne, *Digital Discovery*, 2025, **4**, 2439–2453.
57 U. Schmocker, H. Boesch and F. Waldner, *Phys. Lett. A*, 1972, **40**, 237–238.
58 H. Furuhashi, M. Inagaki and S. Naka, *J. Inorg. Nucl. Chem.*, 1973, **35**, 3009–3014.
59 E. Gorter, *Adv. Phys.*, 1957, **6**, 336–361.
60 S. Carbonin, U. Russo and A. Della Giusta, *Mineral. Mag.*, 1996, **60**, 355–368.
61 H. S. C. O'Neill and A. Navrotsky, *Am. Mineral.*, 1983, **68**, 181–194.
62 D. S. McClure, *J. Phys. Chem. Solids*, 1957, **3**, 311–317.
63 A. Navrotsky and O. Kleppa, *J. Inorg. Nucl. Chem.*, 1967, **29**, 2701–2714.
64 A. Zunger, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 1980, **22**, 5839.
65 X. Zhang and A. Zunger, *Adv. Funct. Mater.*, 2010, **20**, 1944–1952.
66 R. Liu, E. Liu, J. Riebesell, J. Qi, S. P. Ong and T. W. Ko, *MatCalc*, 2024, https://github.com/materialsvirtuallab/matcalc.
67 K. Okada, H. Morikawa, F. Marumo and S. Iwai, *Acta Crystallogr., Sect. B*, 1974, **30**, 1872–1873.
68 S. A. Abbas, I. Mahmood, M. Sajjad, N. Noor, Q. Mahmood, M. Naeem, A. Mahmood and S. M. Ramay, *Chem. Phys.*, 2020, **538**, 110902.
69 A. Togo, L. Chaput, T. Tadano and I. Tanaka, *J. Phys.:Condens. Matter*, 2023, **35**, 353001.
70 A. Togo, *J. Phys. Soc. Jpn.*, 2023, **92**, 012001.