The limited predictive power of the Pauling rules

Janine George, a David Waroquiers, a Davide Di Stefano, a Guido Petretto, a Gian-Marc Rignanese, a and Geoffroy Hautier* a

a Institute of Condensed Matter and Nanosciences, Université Catholique de Louvain, Chemin des étoiles 8, 1348 Louvain-la-Neuve, Belgium, E-mail: geoffroy.hautier@uclouvain.be

Abstract

The Pauling rules have been used for decades to rationalize the crystal structures of ionic compounds. Despite their importance, there has been so far no statistical assessment of the performances of these five empirical rules. Here, we test rigorously and automatically all five Pauling rules for a large data set of around 5000 known oxides. We discuss each Pauling rule separately stressing their limits and range of application in terms of chemistries and structures. We conclude that only 13% of the oxides simultaneously satisfy the last four rules, indicating their much lower predictive power than expected.

Introduction

Understanding and predicting the crystal structure of inorganic materials is an important goal of chemistry. In 1929, Linus Pauling published a series of five empirical rules rationalizing inorganic crystal structures. The Pauling rules apply to ionic compounds and describe what are the preferred local environments of a cation and how these environments connect to each other. These rules have become a cornerstone of solid-state chemistry and remain the main empirical theory rationalizing crystal structure stability. Pauling developed these five rules by combining his knowledge of inorganic crystal structures and simple electrostatic arguments. Though, in response to observed deviations, these rules have been slightly improved over the years, they remain widely used nowadays in their original form. The Pauling rules are not laws of nature. It is thus expected that they are not always correct. However, they have not been yet assessed statistically on a large scale. The absence of such a rigorous assessment of the validity of the Pauling rules inhibits their use for true prediction and prevents the development of improved and alternative rules. Building on recent advances in crystal structure analysis tools including the automatic identification of local environments and their connectivity, we report here on the first statistical evaluation of the Pauling rules on several thousands of compounds. Our work shines light on their strengths and limits, enables their more cautious use and offers a first necessary step towards their future improvement.

Results and Discussion

Our analysis relies on the use of a tool for automatic local environment detection on a set of oxides coming from the Inorganic Crystal Structure Database (ICSD) and present in the Materials Project database (see Methods). We focused on oxides as they are ionic enough for the Pauling rules to be applicable and because the large number of oxides known makes it possible to obtain a large data set and hence good statistics. In total, more than 5000 oxides (a subset of the structures from Ref. 9) are considered. The paper presents and discusses the performance of each of the five rules individually and wraps up by commenting on the overall quality of the five rules taken altogether.

1. Radius Ratio Rule: The first rule describes how the coordination number/coordination environment of a cation is determined by its size: "[...] The coordination number of the cation [is determined] by the radius ratio [of cation and anion].". This rule is based on a hard sphere model of the atoms as shown in
Figure 1a. A coordination environment is stable only if the radius ratio of cation and anion falls within the geometrically derived stability window of this environment.

Figure 1. Assessment of the first rule. a) Illustration of the radius-ratio rule (Pauling’s first rule). Based on purely geometric considerations, it allows one to decide whether a coordination polyhedron is stable. The relevant criterion is the ratio between the radii of the cation and the anion building the coordination polyhedron. The figure was inspired by reference 13. b) Percentages of correct predictions by Pauling’s rule to all tested predictions. The prediction is correct if it agrees with the geometrical assessment of the structure. We only tested those environments where our geometrical assessment resulted in an environment listed in Table S1 in the Supporting Information.

While several atomic and ionic radii schemes were developed after Pauling’s original work, we used here the simplest – Pauling’s univalent radii.7 We found a rather unsatisfactory agreement with the first rule in our data set. Only 66% of the tested local environment agreed with the expectation from Pauling’s first rule. Figure 1b shows an elementwise analysis of the fulfilment of the rule. Very good agreement is observed especially for main group elements (e.g. Si, P and S) but most of the periodic table and especially alkali, alkaline-earth and transition metals present strong deviations from the rule. By design, the first rule can only work for elements presenting a low diversity in local environments (Si, P and S are for instance mainly tetrahedral).9 In contrast, many of the alkali and alkaline-earth metals and some transition metals are found in a variety of environments. We link the failure of the first rule to an inadequate fundamental assumption that a given cation should always be found in one and only one type of local environment. Crystal chemistry is more complex and many cations show diversity in their possible local environments in oxides (see Figure S1 of the Supplementary Information).9 While other ionic radii have been proposed to improve the first rule, we note that some of them such as the Shannon radii6 use ionic radii depending on the local environment and cannot be used for local environment prediction.
2. Electrostatic Valence Rule: The second rule focuses on local charge compensation within crystal structures: "In a stable coordination structure the electric charge of each anion tends to compensate the strength of the electrostatic valence bonds reaching to it from the cations at the centres of the polyhedra of which it forms a corner [...]." This can be thought of as a rule of local charge balance. The application of this rule is demonstrated on α-quartz (see Figure 2a). The charge of the oxygen atom (-2) is compensated by the bond strengths of the two silicon neighbours (+1). The bond strength of the bond from the cation is calculated by dividing the valence of the cation by its coordination number (+4/4).

We tested the second rule for all oxygen atoms in our data set (Figure 2b, left). The rule is only nearly exactly fulfilled for roughly 20% of all oxygen atoms (absolute deviation of 0.01 is allowed). Extreme deviations from the second rule are observed in some common structures such as pyrochlores. When mixing 2+ and 5+ cations (e.g. Cd₂Ta₂O₇, mp-5548) the O in 2a Wyckoff position in pyrochlore shows a sum of bond strengths of one (versus an expected value of 2).

Figure 2c indicates the fulfilment of the second rule per element. There is no clear trend within the periodic table. Oxygens surrounded by Al, Si, and Sn as well as certain transition metals (Sc, Rh, Pd, Ir) satisfy this rule best (~60%). In contrast, oxygens surrounded by P deviate strongly from the rule despite its close proximity to Si in the periodic table. Many phosphates and vanadates are in fact among the exceptions to the second rule, e.g. InPO₄ (Materials Project ID: mp-7566), and CrVO₄ (mp-19418) which crystallise in the same structure type. InPO₄ and the calculation of the bond strengths sums are depicted in Figure 2d. In this simple structure, the oxygens show a sum of bond strengths deviating by 12.5% from the nominal +2 value.

According to Baur, distortions of chemical environments could accommodate deviations from local charge compensation. Our data agrees with this hypothesis and a subset of highly symmetrical, undistorted, local environments (221 materials instead of ~5100) lead to a nearly perfect fulfilment of the rule as shown in Figure 2b (right). In a nutshell, the second rule can be powerful but only for very symmetric structures with undistorted environments. Unfortunately, these perfect structures are a minority of the existing oxides.
Figure 2. Assessment of the second rule. a) Illustration of the electrostatic valence rule (Pauling's second rule). b) Share of the oxygen atoms that show at most a certain absolute deviation from the ideal valence -2 for all environments (left) and for structures with only very symmetric coordination environments (right, the environments show a maximum continuous symmetry measure lower than 0.1 and are assigned with a high certainty, i.e. a coordination environment fraction higher than 0.95, see reference 9 for more information). c) Elementwise depiction of the fulfillment of the second rule. d) Depiction of relevant connections of polyhedra for the calculation of the sum of bond strengths within the crystal structure of InPO$_4$. The sum of bond strengths arrives in one case at 2.25 and in the other at 1.75.

3. The Sharing of Edges and Faces:
The third rule links the stability of crystal structures with the type of connections between the coordination polyhedra as illustrated in Figure 3: “The presence of shared edges, and particularly of shared faces, in a coordinated structure decreases its stability; [...]” Simple electrostatic arguments can justify such a rule as cations will be closer when sharing edges and faces.

**Figure 3.** Assessment of the third rule. a) Illustration of the Rule on “The Sharing of Edges and Faces” (Pauling’s third rule). In stable crystals, corner connections of coordination polyhedra are preferred over edge and especially face connections. b) Shares of connected pairs of polyhedra that are connected via corners, edges and faces for all coordination polyhedra (left) and only coordination polyhedra with a coordination number (CN) equal to 8 or smaller (right). c) Elementwise fulfillment of the rule. The fulfillment of the rule seems to depend on the atomic radius of the elements. d) Structure of BaTiO₃ (mp-5986) and Al₂O₃ (corundum, mp-1143) showing connections of polyhedra via faces.
We tested the third rule by identifying all connected pairs of polyhedra and computing the fraction of the connected pairs that are corner-, edge- or face-sharing (see Figure 3b (left)). The connections via corners dominate with 63%. Then, the edge connections follow with 27%. The lowest percentage of connected polyhedra pairs are connected via faces (10%). This agrees well with Pauling's rule. To probe the effect of chemistry on the third rule, we show in Figure 3c how often face-sharing connections are present for different cations. In general, smaller elements seem to fulfill the rule better (low period and high group number). This is confirmed by a clear dependence of cation connectivity with atomic radius (see SI Fig. S8 in the Supporting Information). Many of the “deviations” to the rule, for instance the presence of face-sharing polyhedra, come from larger cations. For instance, the very common perovskite structure shows face-sharing polyhedra but only for the large cation A site (e.g. Ba$^{2+}$ in BaTiO$_3$). In fact, if we exclude cations with high coordination numbers from our analysis (considering only coordination numbers ≤8), we obtain an even better agreement with Pauling’s third rule with less than two percent of face-sharing environments (see Figure 3b, right). Structures with face-sharing low-coordination ions such as the corundum structure (Al$_2$O$_3$, mp-1143, see Figure 3d)$^{19, 20}$ are quite exceptional. The importance of the coordination number is rationalised by its link to ionic size and to cation-cation distances. Larger ions will have larger cation-cation distances and lower electrostatics.

4. The Nature of Contiguous Polyhedra: The fourth rule is an extension of the third rule. It focuses on crystals containing different cations and indicates how they should connect depending on their oxidation state and local environment (see Figure 4a for a depiction of the rule): “In a crystal containing different cations those with large valence and small coordination number tend not to share polyhedron elements with each other.” Similar to the third rule, we analysed pairs of polyhedra but included non-connected pairs of polyhedra as well. We plot the share of connected pairs of all considered pairs of polyhedra (selected by a distance criterion) as a function of the coordination numbers and as a function of the valences (see Figure 4b). We find that the coordination number strongly affects the tendency for two polyhedra to be connected. Cations with lower coordination numbers tend to clearly be less connected to each other. This is in good agreement with Pauling fourth rule. However, the oxidation state surprisingly does not seem to affect the connectivity between polyhedra in clear disagreement with the fourth rule. The exceptions to the fourth rule (i.e. structures in which the polyhedra of cations with the highest valence and smallest coordination number are connected) amount to 40% of all tested structures. Many structures mixing octahedral sites filled with a high valent cation and higher coordination number sites with low valent cations do not fulfil the fourth rule. This is the case for instance for the perovskite (e.g. BaTiO$_3$ (mp-5986))$^{17, 18}$, La$_3$Nb$_2$O$_7$ (mp-560349)$^{21}$ and pyrochlore structures (e.g. Nb$_2$Cd$_2$O$_7$ (mp-5472)).$^{22}$ In all these examples, the high valent (4+ or 5+) octahedra environments are directly connected together in direct violation of the fourth rule. This indicates that covalent and electronic structure effects might be more important than electrostatics in these compounds. Additionally, structures belonging to the α-NaFeO$_2$ structure type$^{23}$ (mp-19359, see Figure 4d for a depiction of the structure) are also among the exceptions. Here, all cations are octahedrally coordinated and the cations with the largest valence share polyhedron elements with each other. In the specific case of α-NaFeO$_2$, this deviation from the Pauling rules might be rationalised with the help of antiferromagnetic interactions of the iron atoms$^{24, 25}$ which might stabilise the structure only when the Fe ions are direct neighbours.
Figure 4. Assessment of the fourth rule. a) Depiction of the fourth rule which deals with “The Nature of Contiguous Polyhedra”: a stable crystal should not exhibit connections between coordination polyhedra with a small coordination number (CN) and around a cation with a large valence. b) The dependency of the share of connected pairs of polyhedra on the coordination numbers, and valences of the cations. c) Elementwise depiction of the rule fulfilment. d) One of the exceptions, α-NaFeO₂, is shown. The octahedral coordination environments of Fe(III) show connections. Na⁺ is also octahedrally coordinated so that Fe(III) is the cation with the lowest coordination number and the highest valence.
Figure 4c shows the dependence of the fourth rule on chemistry. Main group elements in general follow the rule more than other chemistries such as transition metals. Within the main group elements, we see however strong differences. For instance, sulfur (in SO$_4^{2-}$ groups) almost always fulfils the rule, while this is less the case for phosphorus. This comes from the less common condensation of polyanions in sulfates compared to phosphates with for instance 128 structures forming P$_2$O$_7^{4-}$ groups while the S$_2$O$_7^{2-}$ groups are rare (5 within our database).

5. The Rule of Parsimony: The fifth rule states the number of constituents in stable crystal structures. According to Pauling, it should hold that “The number of essentially different kinds of constituents in a crystal tends to be small.” In other words, a cation (an element in a given oxidation state) prefers to occupy the same local environment in a given crystal structure. For instance, it would be unfavourable for a cation to be tetrahedral and octahedral in the same compound (see Figure 5a).

![Rule of Parsimony Diagram]

**Figure 5.** Assessment of the fifth rule. a) Visualisation of the Rule of Parsimony (Pauling’s fifth rule). Crystals with unit cells that only include one kind of coordination environment of chemically identical cations are more stable than others including more than one kind of coordination environment. b) Pie
chart depicting the structures fulfilling the rule of parsimony (only coordination numbers are considered).

c) Exceptions: Crystal structure of $\beta$-Ga$_2$O$_3$\textsuperscript{26,27} and As$_2$O$_5$\textsuperscript{28}. The crystal structure of $\beta$-Ga$_2$O$_3$\textsuperscript{26,27} shows both tetrahedral and octahedral coordination of Ga(+III). $\beta$-Ga$_2$O$_3$ is more stable than $\alpha$-Ga$_2$O$_3$ (energy above hull and ref. 26). The structure of As$_2$O$_5$ shows both tetrahedral and octahedral environments for As(+V).

d) Elementwise fulfilment of the fifth rule when only coordination numbers are considered.

Around 70% of all tested structures show the same local environment for their cation types (see Figure 5b). The elementwise analysis of the rule fulfilment is depicted in Fig 5c. Alkali and alkaline-earth metals deviate from this rule and are easily present in different local environments in the same structure. The wide range of possible local environment for alkali and alkaline-earth metals makes the breaking of the fifth rule not surprising. More surprisingly, a few main group elements such as B, Ga or Ge also break quite commonly the parsimony rule. Borates mix easily tetrahedral and trigonal planar borate groups such as in CaB$_2$O$_4$ (mp-8056)\textsuperscript{29}. Exceptions to the rule are even found in simple binaries: $\beta$-Ga$_2$O$_3$ (mp-886)\textsuperscript{26,27} and As$_2$O$_5$ (mp-1788) contain Ga(+III) or As(+V) in fourfold and sixfold coordination within the same compound (see Figure 5c).\textsuperscript{28}

**Combined Assessment of the Five Rules:** To assess the overall performance of the Pauling rules, we combine the assessment of the second to fifth rules. We avoid assessing the first rule, because it will strongly depend on the quality of the univalent radii. Additionally, we do not believe that any simple rule directly linking ionic radius to local environment will ever be predictive in view of the local environment diversity of many cations.
We found a few structure prototypes that fulfil the 4 Pauling rules: the rutile structure type such as SnO$_2$ (mp-856), spinels such as MgAl$_2$O$_4^{30,31}$ (mp-3536), structures in the scheelite structure type (CaWO$_4$) such as BiAsO$_3^{32}$ (mp-561068), and many structures in the ZrSiO$_4$ (mp-4820) structure type. Furthermore, some phosphates such as $\alpha$-AlPO$_4$ (mp-3955, berlineite) which is isotypic to $\alpha$-quartz,$^{33,34}$ several sulfates such as In$_2$(SO$_4$)$_3$ (mp-541450),$^{35}$ and silicates such as Mg$_2$SiO$_4$ (mp-2895)$^{36}$ are among the fulfilling structures. If one looks at the elementwise fulfilment of these four rules in Figure 6b, one can see that structures including Al, and Si tend to fulfil the rule better than structures including P and the alkali and alkaline-earth metals.

Overall, only 13% of all tested structures fulfil the second to fifth rules simultaneously. This indicates that structures with at least one deviation to the Pauling rules are frequent and demonstrates that the predictive power of the Pauling rules is quite limited. To find out what rules are the most problematic, we show in Figure 6a how our conclusion varies if some rules are disregarded. Each green bar indicating the success rate of the Pauling rules if the 4 rules are combined or only a subset of three rules. The second rule leads to the most exceptions and has the largest influence on the overall fulfilment of the rules. As the previous analysis of the rules indicated the different behaviour of high versus low coordination environments, we plot in Figure 6a in blue the fulfilment of the rule when excluding cations in high coordination environment. Even in those more favourable conditions, the improvement is modest with 20% of compounds fulfilling the four rules.

**Conclusion**

We presented the first statistical analysis of all five Pauling rules focusing on their individual performances and their dependency on chemistry. In our opinion, the first rule, while commonly used, is the most problematic. Indeed, there are many cations that show a large diversity of local environments and fundamentally cannot be correctly described by the first rule. Instead of using the radius-ratio rule, we suggest using statistics of the coordination environments to check if the coordination environment found is usual or not.

As for the other rules, only 13% of all tested structures fulfil the second to fifth rules simultaneously. Our results cast doubt on the real predictive power of the Pauling rules. The success rate of the rules can be improved by narrowing their application to certain regions of the periodic table or to certain cations. Restricting the rules to certain chemical families does not appear very promising as no clear chemical trend across all rules is observed and elements as chemically similar as Si and P obey the rules very differently. Ionic size appears more important and removing large coordination environments (>8) from the analysis improves significantly the third and fourth rules. Likewise, only considering undistorted, very symmetric, local environment improves dramatically the second rule. Our findings about the true scope of different Pauling rules will be very helpful to the solid-state chemists commonly applying them. However, restricting the rules to specific cases lowers the universal ambition of the original Pauling rules as the vast majority of the oxides lie outside these restrictions. Our work therefore calls for the development of new empirical rules beyond the almost one-century old Pauling rules. Our analysis and the data set of connectivity and local environment provided is a first step towards building this new theory that could potentially benefit from the recent growth in the use of machine learning techniques in chemistry and materials science.$^{37,38}$
Methods:

We analysed the \textit{ab initio} relaxed structures of around 5100 oxides from the Materials Project Database that originally stem from the experimental part of the Inorganic Crystal Structure Database (ICSD)\textsuperscript{39} to ensure that no pathologically unstable structures are included in our analysis. We only focused on materials with an energy above hull smaller or equal to 0.025 eV/atom (status on 10.4.2019). This \textit{ab initio} relaxation corrects for uncertainties of the bond lengths and especially the coordination environments within the original structure predictions via X-ray or neutron diffraction. We arrive at quantitatively similar results if we use purely experimental data of 5000 oxides from the Crystallography Open Database (Figures S2-S7 in the Supporting Information).\textsuperscript{40, 41, 42, 43, 44} We ensured that no duplicates were present in this data set and we worked with the primitive cells. The oxidation states are determined with the help of a bond-valence analysis that is implemented in pymatgen.\textsuperscript{45} If this analysis failed, we used the oxidation states as given in the ICSD (only for the Materials Project data).\textsuperscript{9} To analyse the coordination environments, we used the ChemEnv package with the default MultiWeightsChemEnvStrategy, that was developed in the previous study on the statistics of coordination environments in oxides.\textsuperscript{9} We applied it to a primitive unit cell of the oxide. To analyse the Pauling rules, we then only used the coordination environment that was assigned with the highest confidence (i.e. highest “ce\_fraction”). The oxides from the Materials Project have already been studied in reference \textsuperscript{9} and the assignment of the oxidation states was based on this study. The code to analyse each of the Pauling rules can be accessed via https://github.com/jageo/paulingpublication and it is described in detail in the Supporting Information. CSV Files including the structures fulfilling and not fulfilling the second to fifth rules are included in the Supporting Information.

All crystal structures have been depicted with VESTA.\textsuperscript{46}

Acknowledgements

We would like to thank all crystallographers and inorganic chemists filling the crystal structure databases with entries. We acknowledge the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CÉCil) for computational resources. J. G. acknowledges funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 837910.

Author Contributions

G. H. and J. G. designed the study with inputs from all authors. J. G. wrote the code to analyze the Pauling Rules with inputs from D. W., D. D.S, G. P. G.H. and G.-M. R. managed the project. All authors analyzed and discussed the results. G. H. and J. G. wrote the manuscript and it was approved by all authors.

Competing Interests statement

The authors declare no competing interests.
References

1. Pauling L. The principles determining the structure of complex ionic crystals. *J Am Chem Soc* 1929, 51(4): 1010.

2. Baur WH. Bond length variation and distorted coordination polyhedra in inorganic crystals. *Trans Am Crystallogr Assoc* 1970, 6: 129-155.

3. Brese N, O'Keeffe M. Bond-Valence Parameters for Solids. *Acta Cryst B* 1991, 47(2): 192–197.

4. O'Keeffe M, Brese NE. Atom sizes and bond lengths in molecules and crystals. *J Am Chem Soc* 1991, 113(9): 3226-3229.

5. Beck HP. A study on AB2O6 compounds, Part III: Co-ordination needs and patterns of connectivity. *Z Kristallogr - Cryst Mater* 2013, 228(6): 271.

6. Shannon R. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Cryst A* 1976, 32(5): 751.

7. Pauling L. *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*. Cornell University Press, 1960.

8. Burdett JK, Price SL, Price GD. The factors influencing solid state structure — A modern Mendeleyevian answer. *Solid State Communications* 1981, 40(10): 923-926.

9. Waroquiers D, Gonze X, Rignanese G-M, Welker-Nieuwoudt C, Rosowski F, Gölbe M, et al. Statistical Analysis of Coordination Environments in Oxides. *Chem Mat* 2017, 29(19): 8346.

10. Zimmermann NER, Horton MK, Jain A, Haranczyk M. Assessing Local Structure Motifs Using Order Parameters for Motif Recognition, Interstitial Identification, and Diffusion Path Characterization. *Front Mater* 2017, 4: 34.

11. Zimmermann N, Jain A. Local Structure Order Parameters and Site Fingerprints for Quantification of Coordination Environment and Crystal Structure Similarity. 2019.

12. Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, et al. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater* 2013, 1(1): 011002.

13. Peng W, Petretto G, Rignanese G-M, Hautier G, Zevalkink A. An Unlikely Route to Low Lattice Thermal Conductivity: Small Atoms in a Simple Layered Structure. *Joule* 2018, 2(9): 1879-1893.

14. Bystroem A. Calcium pyroantimonates and similar compounds. *Ark Kemi, Mineral Geol* 1945, 18: 8.

15. Mooney R. Crystal structure of anhydrous indium phosphate and thallic phosphate by X-ray diffraction. *Acta Cryst* 1956, 9(2): 113-117.

16. Frazer BC, Brown PJ. Antiferromagnetic Structure of CrVO4 and the Anhydrous Sulfates of Divalent Fe, Ni, and Co. *Phys Rev* 1962, 125(4): 1283-1291.

17. Evans HT. The crystal structure of tetragonal barium titanate. *Acta Cryst* 1951, 4(4): 377.
18. Megaw H. Refinement of the structure of BaTiO3 and other ferroelectrics. *Acta Cryst* 1962, **15**(10): 972-973.

19. Bragg WH. The significance of crystal structure. *J Chem Soc, Trans* 1922, **121**(0): 2766-2787.

20. Pauling L, Hendricks SB. The crystal structures of hematite and corundum. *J Am Chem Soc* 1925, **47**(3): 781-790.

21. Lam R, Wiss F, Greedan JE. Magnetic Properties of the Fluorite-Related La3MO7 Phases, M=Ru and Os: Local Moment Magnetism, Short- and Long-Range Order in 4d and 5d Transition Metal Oxides. *J Solid State Chem* 2002, **167**(1): 182-187.

22. Brisse F, Stewart DJ, Seidl V, Knop O. Pyrochlores. VIII. Studies of some 2–5 Pyrochlores and Related Compounds and Minerals. *Can J Chem* 1972, **50**(22): 3648-3666.

23. Goldsztäub S. Étude de quelques dérivés de l’oxyde ferrique (FeO,OH, FeO2Na,FeOCI) ; détermination de leurs structures. *Bull Minéral* 1935: 6-76.

24. Tomkowicz Z, van Laar B. Magnetic properties and magnetic structure of α-NaFeO2. *Phys Status Solidi A* 1974, **23**(2): 683-688.

25. McQueen T, Huang Q, Lynn JW, Berger RF, Klimczuk T, Ueland BG, et al. Magnetic structure and properties of the S=52 triangular antiferromagnet alpha-NaFeO2. *Phys Rev B* 2007, **76**(2): 024420.

26. Geller S. Crystal Structure of β-Ga2O3. *J Chem Phys* 1960, **33**(3): 676-684.

27. Ahnman J, Svensson G, Albertsson J. A Reinvestigation of [beta]-Gallium Oxide. *Acta Cryst C* 1996, **52**(6): 1336-1338.

28. Jansen M. Crystal Structure of As2O5. *Angew Chem Int Ed* 1977, **16**(5): 314-315.

29. Shashkin D, Simonov M, Belov NVe. Crystalline structure of calciborite—CaB_2O_4=Ca_2BO_3BO_2. Doklady Akademii Nauk; 1970: Russian Academy of Sciences; 1970. p. 345-348.

30. Bragg WH. The Structure of Magnetite and the Spinels. *Nature* 1915, **95**(2386): 561-561.

31. Passerini L. Ricerche sugli spinelli.—I. I composti CuAl 2 O 4, MgAl 2 O 4, MgFe 2 O 4, ZnAl 2 O 4, ZnCr 2 O 4, ZnFe 2 O 4, MnFe 2 O 4. *Gazz Chim Ital* 1929, **59**: 280-288.

32. Mooney R. Crystal structure of tetragonal bismuth arsenate, BiAsO4. *Acta Cryst* 1948, **1**(4): 163-165.

33. Muraoka Y, Kihara K. The temperature dependence of the crystal structure of berlinité, a quartz-type form of AlPO4. *Phys Chem Miner* 1997, **24**(4): 243-253.

34. Schwarzenbach D. Die Kristallstruktur von AlPO4 (Tiefquarzmodifikation). *Naturwissenschaften* 1965, **52**(12): 343-344.

35. Krause M, Gruehn R. Contributions on the thermal behaviour of sulphates XVII.1 Single crystal structure refinements of In2 (SO4)3 and Ga2 (SO4)3. *Z Kristallog* 1995, **210**(6): 427.

36. Bragg WL, Brown GB. XXX. Die Struktur des Olivins. *Z Kristallog - Cryst Mater* 1926, **63**(1-6): 538.
37. Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. Nature 2018, 559(7715): 547.

38. Schmidt J, Marques MRG, Botti S, Marques MAL. Recent advances and applications of machine learning in solid-state materials science. npj Computational Materials 2019, 5(1): 83.

39. Allmann R, Hinek R. The introduction of structure types into the Inorganic Crystal Structure Database ICSD. Acta Cryst A 2007, 63(5): 412-417.

40. Downs RT, Hall-Wallace M. The American Mineralogist Crystal Structure Database. Am Mineral 2003, 88: 247-250.

41. Grazulis S, Chateigner D, Downs RT, Yokochi AFT, Quiros M, Lutterotti L, et al. Crystallography Open Database - an open-access collection of crystal structures. J Appl Crystallogr 2009, 42(4): 726-729.

42. Gražulis S, Daškevič A, Merkys A, Chateigner D, Lutterotti L, Quiró M, et al. Crystallography Open Database (COD): an open-access collection of crystal structures and platform for worldwide collaboration. Nucleic Acids Res 2012, 40(D1): D420-D427.

43. Gražulis S, Merkys A, Vaitkus A, Okulič-Kazarinas M. Computing stoichiometric molecular composition from crystal structures. J Appl Crystallogr 2015, 48(1): 85-91.

44. Merkys A, Vaitkus A, Butkus J, Okulič-Kazarinas M, Kairys V, Gražulis S. COD::CIF::Parser: an error-correcting CIF parser for the Perl language. J Appl Cryst 2016, 49(1).

45. Ong SP, Richards WD, Jain A, Hautier G, Kocher M, Cholia S, et al. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. Comp Mater Sci 2013, 68: 314.

46. Momma K, Izumi F. VESTA3 for Three-Dimensional Visualization of Crystal, Volumetric and Morphology Data. J Appl Crystallogr 2011, 44(6): 1272–1276.